



Cite this article: Sutter D, Fawzi O, Renner R. 2016 Universal recovery map for approximate Markov chains. *Proc. R. Soc. A* **472**: 20150623. <http://dx.doi.org/10.1098/rspa.2015.0623>

Received: 8 September 2015

Accepted: 24 December 2015

Subject Areas:

quantum physics, mathematical physics,
quantum computing

Keywords:

conditional mutual information,
quantum Markov chains, recoverability,
strong subadditivity

Author for correspondence:

David Sutter

e-mail: sutterd@itp.phys.ethz.ch

Electronic supplementary material is available at <http://dx.doi.org/10.1098/rspa.2015.0623> or via <http://rspa.royalsocietypublishing.org>.

Universal recovery map for approximate Markov chains

David Sutter¹, Omar Fawzi^{2,3} and Renato Renner¹

¹Institute for Theoretical Physics, ETH Zurich, Zurich, Switzerland

²Department of Computing and Mathematical Sciences, Caltech, Pasadena, CA, USA

³LIP, ENS de Lyon, Lyon, France

A central question in quantum information theory is to determine how well lost information can be reconstructed. Crucially, the corresponding recovery operation should perform well without knowing the information to be reconstructed. In this work, we show that the *quantum conditional mutual information* measures the performance of such recovery operations. More precisely, we prove that the conditional mutual information $I(A : C|B)$ of a tripartite quantum state ρ_{ABC} can be bounded from below by its distance to the closest recovered state $\mathcal{R}_{B \rightarrow BC}(\rho_{AB})$, where the C -part is reconstructed from the B -part only and the recovery map $\mathcal{R}_{B \rightarrow BC}$ merely depends on ρ_{BC} . One particular application of this result implies the equivalence between two different approaches to define topological order in quantum systems.

1. Introduction

A state ρ_{ABC} on a tripartite quantum system $A \otimes B \otimes C$ forms a (*quantum*) *Markov chain* if it can be recovered from its marginal ρ_{AB} on $A \otimes B$ by a quantum operation $\mathcal{R}_{B \rightarrow BC}$ from B to $B \otimes C$, i.e.

$$\rho_{ABC} = \mathcal{R}_{B \rightarrow BC}(\rho_{AB}). \quad (1.1)$$

An equivalent characterization of ρ_{ABC} being a quantum Markov chain is that the *conditional mutual information* $I(A : C|B)_\rho := H(AB)_\rho + H(BC)_\rho - H(B)_\rho - H(ABC)_\rho$ is zero [1,2], where $H(A)_\rho := -\text{tr}(\rho_A \log_2 \rho_A)$ is the von Neumann entropy. The structure of these states has been studied in various works. In particular, it has been shown that A and C can be viewed as independent conditioned on B , for a meaningful notion of conditioning [3].

Very recently, it has been shown that Markov states can be alternatively characterized by having a generalized Rényi conditional mutual information that vanishes [4].

A natural question that is relevant for applications is whether the above statements are robust. (In [5] an example is discussed that illustrates why this question is relevant. In [6] further explanations are given to emphasize the importance of this problem.) Specifically, one would like to have a characterization of the states that have a small (but not necessarily vanishing) conditional mutual information, i.e. $I(A : C|B) \leq \varepsilon$ for $\varepsilon > 0$. First results revealed that such states can have a large distance to Markov chains that is independent of ε [7,8], which has been taken as an indication that their characterization may be difficult. However, it has subsequently been realized that a more appropriate measure instead of the distance to a (perfect) Markov chain is to consider how well (1.1) is satisfied [5,9–11]. This motivated the definition of *approximate Markov chains* as states where (1.1) approximately holds.

In recent work [6], it has been shown that the set of approximate Markov chains indeed coincides with the set of states with small conditional mutual information. In particular, the distance between the two terms in (1.1), which may be measured in terms of their fidelity F , is bounded by the conditional mutual information.¹ More precisely, for any state ρ_{ABC} there exists a trace-preserving completely positive map $\mathcal{R}_{B \rightarrow BC}$ (the *recovery map*) such that

$$I(A : C|B)_\rho \geq -2 \log_2 F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})). \quad (1.2)$$

Furthermore, a converse inequality of the form $I(A : C|B)_\rho^2 \leq -c^2 \log_2 F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))$, where c depends logarithmically on the dimension of A can be shown to hold [6,11].

We also note that the fidelity term in (1.2), maximized over all recovery maps, i.e.

$$F(A; C|B)_\rho := \sup_{\mathcal{R}_{B \rightarrow BC}} F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) \quad (1.3)$$

is called *fidelity of recovery*² and has been introduced and studied in [14,15]. With this quantity, the main result of [6] can be written as

$$I(A : C|B)_\rho \geq -2 \log_2 F(A; C|B)_\rho. \quad (1.4)$$

The fidelity of recovery has several natural properties, e.g. it is monotonous under local operations on A and C , and it is multiplicative [15].

The result of [6] has been extended in various ways. Based on quantum state redistribution protocols, it has been shown in [16] that (1.2) still holds if the fidelity term is replaced by the *measured relative entropy* $D_{\mathbb{M}}(\cdot, \cdot)$, which is generally larger, i.e. there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that

$$I(A : C|B)_\rho \geq D_{\mathbb{M}}(\rho_{ABC} \| \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) \geq -2 \log_2 F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})). \quad (1.5)$$

The measured relative entropy is defined as the supremum of the relative entropy with measured inputs over all projective measurements³ $\mathcal{M} = \{M_x\}$, i.e.

$$D_{\mathbb{M}}(\rho \| \sigma) := \sup \left\{ D(\mathcal{M}(\rho) \| \mathcal{M}(\sigma)) : \mathcal{M}(\rho) = \sum_x \text{tr}(\rho M_x) |x\rangle \langle x| \text{ with } \sum_x M_x = \text{id} \right\}, \quad (1.6)$$

where $\{|x\rangle\}$ is a finite set of orthonormal vectors. This quantity was studied in [17,18].

Furthermore, in [15] an alternative proof of (1.2) has been derived that uses properties of the fidelity of recovery (in particular, multiplicativity). Another recent work [19] showed how to generalize ideas from [6] to prove a remainder term for the monotonicity of the relative entropy in terms of a recovery map that satisfies (1.2).

¹The fidelity of ρ and σ is defined as $F(\rho, \sigma) := \|\sqrt{\rho} \sqrt{\sigma}\|_1$.

²We note that if A , B and C are finite-dimensional Hilbert spaces the supremum is achieved, since the set of recovery maps is compact (see remark 10.3 in the electronic supplementary material) and the fidelity is continuous in the input state (see [12,13] or lemma B.9 in [6]).

³Without loss of generality, these can be assumed to be rank-one projectors.

All known proofs of (1.2) are non-constructive, in the sense that the recovery map $\mathcal{R}_{B \rightarrow BC}$ is not given explicitly. It is merely known [6] that if A , B and C are finite-dimensional then $\mathcal{R}_{B \rightarrow BC}$ can always be chosen such that it has the form

$$X_B \mapsto V_{BC} \rho_{BC}^{1/2} \left(\rho_B^{-1/2} U_B X_B U_B^\dagger \rho_B^{-1/2} \otimes \text{id}_C \right) \rho_{BC}^{1/2} V_{BC}^\dagger \quad (1.7)$$

on the support of ρ_B , where U_B and V_{BC} are unitaries on B and $B \otimes C$, respectively. It would be natural to expect that the choice of the recovery map that satisfies (1.2) only depends on ρ_{BC} , however this is only known in special cases. One such special case is Markov chains ρ_{ABC} , i.e. states for which (1.1) holds perfectly. Here a map of the form (1.7) with $V_{BC} = \text{id}_{BC}$ and $U_B = \text{id}_B$ (sometimes referred to as *transpose map* or *Petz recovery map*) serves as a perfect recovery map [1,2]. Another case where a recovery map that only depends on ρ_{BC} is known explicitly are states with a classical B system, i.e. qcq-states of the form $\rho_{ABC} = \sum_b P_B(b) |b\rangle\langle b| \otimes \rho_{AC,b}$, where P_B is a probability distribution, $\{|b\rangle\}_b$ an orthonormal basis on B and $\{\rho_{AC,b}\}_b$ a family of states on $A \otimes C$. As discussed in [6], for such states (1.2) holds for the recovery map defined by $\mathcal{R}_{B \rightarrow BC}(|b\rangle\langle b|) = |b\rangle\langle b| \otimes \rho_{C,b}$ for all b , where $\rho_{C,b} = \text{tr}_A(\rho_{AC,b})$. For general states, however, the previous results left open the possibility that the recovery map $\mathcal{R}_{B \rightarrow BC}$ depends on the *full state* ρ_{ABC} rather than the marginal ρ_{BC} only. In particular, the unitaries U_B and V_{BC} in (1.7), although acting only on B respectively $B \otimes C$, could have such a dependence.

In this work, we show that for any state ρ_{BC} on $B \otimes C$ there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ that is *universal*—in the sense that the distance between *any* extension ρ_{ABC} of ρ_{BC} and $\mathcal{R}_{B \rightarrow BC}(\rho_{AB})$ is bounded from above by the conditional mutual information $I(A : C|B)_\rho$. In other words, we show that (1.2) remains valid if the recovery map is chosen depending on ρ_{BC} only, rather than on ρ_{ABC} . This result implies a close connection between two different approaches to define topological order of quantum systems.

2. Main result

Theorem 2.1. *For any density operator ρ_{BC} on $B \otimes C$, there exists a trace-preserving completely positive map $\mathcal{R}_{B \rightarrow BC}$ such that for any extension ρ_{ABC} on $A \otimes B \otimes C$*

$$I(A : C|B)_\rho \geq -2 \log_2 F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})), \quad (2.1)$$

where A , B and C are separable Hilbert spaces.

Remark 2.2. If B and C are finite-dimensional Hilbert spaces, the statement of theorem 2.1 can be tightened to

$$I(A : C|B)_\rho \geq D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB})). \quad (2.2)$$

Remark 2.3. The recovery map $\mathcal{R}_{B \rightarrow BC}$ predicted by theorem 2.1 has the property that it maps ρ_B to ρ_{BC} . To see this, note that $I(A : C|B)_{\tilde{\rho}} = 0$ for any density operator of the form $\tilde{\rho}_{ABC} = \rho_A \otimes \rho_{BC}$. Theorem 2.1 thus asserts that $\tilde{\rho}_{ABC}$ must be equal to $\mathcal{R}_{B \rightarrow BC}(\tilde{\rho}_{AB})$, which implies that $\rho_{BC} = \mathcal{R}_{B \rightarrow BC}(\rho_B)$. We note that so far it was unknown whether recovery maps that satisfy (1.2) and have this property do exist.

We note that theorem 2.1 does not reveal any information about the structure of the recovery map that satisfies (2.1). However, if we consider a linearized version of the bound (2.1), we can make more specific statements.

Corollary 2.4. *For any density operator ρ_{BC} on $B \otimes C$, there exists a trace-preserving completely positive map $\mathcal{R}_{B \rightarrow BC}$ such that for any extension ρ_{ABC} on $A \otimes B \otimes C$*

$$I(A : C|B)_\rho \geq \frac{2}{\ln(2)} (1 - F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))) \quad (2.3)$$

where A , B and C are separable Hilbert spaces. Furthermore, if B and C are finite-dimensional then $\mathcal{R}_{B \rightarrow BC}$ has the form

$$X_B \mapsto \rho_{BC}^{1/2} \mathcal{U}_{BC \rightarrow BC} (\rho_B^{-1/2} X_B \rho_B^{-1/2} \otimes \text{id}_C) \rho_{BC}^{1/2} \quad (2.4)$$

on the support of ρ_B , where $\mathcal{U}_{BC \rightarrow BC}$ is a unital trace-preserving map from $B \otimes C$ to $B \otimes C$.

Remark 2.5. Following the proof of corollary 2.4, we can deduce a more specific structure of the universal recovery map. In the finite-dimensional case, the map $\mathcal{R}_{B \rightarrow BC}$ satisfying (2.3) can be assumed to have the form

$$X_B \mapsto \int V_{BC}^s \rho_{BC}^{1/2} (\rho_B^{-1/2} U_B^s X_B U_B^{s\dagger} \rho_B^{-1/2} \otimes \text{id}_C) \rho_{BC}^{1/2} V_{BC}^{s\dagger} \mu(ds), \quad (2.5)$$

where μ is a probability measure on some set \mathcal{S} , $\{V_{BC}^s\}_{s \in \mathcal{S}}$ is a family of unitaries on $B \otimes C$ that commute with ρ_{BC} , and $\{U_B^s\}_{s \in \mathcal{S}}$ is a family of unitaries on B that commute with ρ_B . However, the representation of the recovery map given in (2.4) has certain advantages compared to the representation (2.5). The fidelity maximized over all recovery maps of the form (2.4) can be phrased as a semidefinite programme and therefore be computed efficiently, whereas it is unknown whether the same is possible for (2.5).

We note that for almost all density operators ρ_{BC} , i.e. for all ρ_{BC} except for a set of measure zero, we can replace the unitaries U_B^s and V_{BC}^s by complex matrix exponentials of the form ρ_B^{it} and ρ_{BC}^{it} , respectively, with $t \in \mathbb{R}$. This shows that (2.5) without the integral (the integration in (2.5) is only necessary to guarantee that the recovery map is universal) coincides with the recovery map found in [20].⁴

Example 2.6. For density operators with a marginal on $B \otimes C$ of the form $\rho_{BC} = \rho_B \otimes \rho_C$, a universal recovery map that satisfies (2.2) is uniquely defined on the support of ρ_B —it is the transpose map, which in this case simplifies to $\mathcal{R}_{B \rightarrow BC} : X_B \mapsto X_B \otimes \rho_C$. It is straightforward to see that (2.2) holds. In fact, we even have equality if we consider the relative entropy (which is in general larger than the measured relative entropy), i.e.

$$I(A : C|B)_\rho = D(\rho_{ABC} \| \mathcal{R}_{B \rightarrow BC}(\rho_{AB})). \quad (2.6)$$

The uniqueness of $\mathcal{R}_{B \rightarrow BC}$ on the support of ρ_B follows by using the fact that the universal recovery map should perfectly recover the Markov state $\rho_{AB} \otimes \rho_C$ where ρ_{AB} is a purification of ρ_B . This forces $\mathcal{R}_{B \rightarrow BC}$ to agree with the transpose map on the support of ρ_B [1,2].

The proof of theorem 2.1 is structured into two parts. We first prove the statement for finite-dimensional Hilbert spaces B , and C in §4 and then show that this implies the statement for general separable Hilbert spaces in §5. The proof of corollary 2.4 is given in §6.

3. Applications

A celebrated result known as *strong subadditivity* states that the conditional quantum mutual information of any density operator is non-negative [23,24], i.e.

$$I(A : C|B)_\rho \geq 0, \quad (3.1)$$

for any density operator ρ_{ABC} on $A \otimes B \otimes C$. Theorem 2.1 implies a strengthened version of this inequality with a remainder term that is universal in the sense that it only depends on ρ_{BC} . The conditional quantum mutual information is a useful tool in different areas of physics and computer science. It is helpful to characterize measures of entanglement [6,25], analyse the correlations of quantum many-body systems [5,26], prove quantum de Finetti results [27,28] and make statements about quantum information complexity [29–31]. It is expected that oftentimes when (1.2) can be used, its universal version (predicted by theorem 2.1) is even more helpful.

In the following, we sketch an application where the universality result is indispensable. Theorem 2.1 can be applied to establish a connection between two alternative definitions of

⁴This follows by the equidistribution theorem which is a special case of the strong ergodic theorem [21, §II.5] (see also [22]).

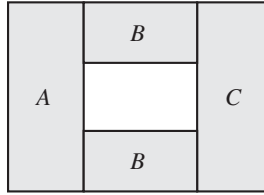


Figure 1. Relevant topology of the subsystems A, B and C such that a state ρ_{ABC} exhibits TQO' if $I(A : C|B)_\rho = 2\gamma > 0$.

topological order of quantum systems (denoted by TQO and TQO'). Consider an n -spin system with $n \in \mathbb{N}$. While the following statements should be understood asymptotically (in the limit $n \rightarrow \infty$), we omit the dependence on n in our notation for simplicity.

According to [32], a family of states $\{\rho^i\}_{i \in \mathcal{I}}$ with $\rho^i \in \mathcal{E}$ for all $i \in \mathcal{I}$ and $|\mathcal{I}| < \infty$ where \mathcal{E} denotes a collection of states, exhibits topological quantum order (TQO) if and only if any two members of the family:

- (i) are (asymptotically) orthogonal, i.e. $F(\rho^i, \rho^j) = 0$ for all $i \neq j \in \mathcal{I}$ and
- (ii) have (asymptotically) the same marginals on any sufficiently small subregion, i.e. $\text{tr}_G \rho^i = \text{tr}_G \rho^j$ for all $i, j \in \mathcal{I}$ and G sufficiently large.⁵

Alternatively, for three regions A, B and C that form a certain topology \mathcal{F} (see figure 1 and [33]), a state ρ_{ABC} on such a subspace exhibits topological quantum order (TQO') if $I(A : C|B)_\rho = 2\gamma > 0$, where γ denotes a *topological entanglement entropy* [33].⁶ (See [33] for more explanations on how the topological entanglement entropy is defined for the topology \mathcal{F} depicted in figure 1.)

It is an open problem to find out how these two characterizations are related, e.g. if a family \mathcal{K} of states on \mathcal{F} that exhibits TQO implies that most of its members have TQO'. This connection follows by theorem 2.1. Suppose $\{\rho^i\}_{i \in \mathcal{I}}$ with $\rho^i \in \mathcal{K}$ for all $i \in \mathcal{I}$ shows TQO. Then consider subsystems A, B and C that together form a non-contractible loop. By definition of TQO, the density operators $\{\rho^i\}_{i \in \mathcal{I}}$ share (asymptotically) the same marginals on $B \otimes C$. Applying theorem 2.1 to this common marginal, together with the continuity of the conditional mutual information [35] ensures that there exist a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that for any $i \in \mathcal{I}$,

$$I(A : C|B)_{\rho^i} \geq -2 \log F(\rho_{ABC}^i, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^i)). \quad (3.2)$$

Since the density operators $\{\rho_{ABC}^i\}_{i \in \mathcal{I}}$ are (asymptotically) orthogonal, share (asymptotically) the same marginals on $A \otimes B$, and the fidelity is continuous in its inputs [12,13], this implies that for all $i \in \mathcal{I}$, except of a single element, we have

$$I(A : C|B)_{\rho^i} \geq \text{const.} > 0. \quad (3.3)$$

4. Proof for finite dimensions

Throughout this section, we assume that the Hilbert spaces B and C are finite-dimensional. In the proof Steps 1–3 below, we also make the same assumption for A , but then drop it in Step 4. We start by explaining why (2.2) is a tightened version of (2.1) which was noticed in [16]. Let $D_\alpha(\cdot \| \cdot)$ be the α -Quantum Rényi Divergence as defined in [36,37] with $D_1(\rho \| \sigma) = D(\rho \| \sigma) := \text{tr}(\rho(\log \rho - \log \sigma))$ and $D_\alpha(\rho \| \sigma) := (1/(\alpha - 1)) \log \text{tr}((\sigma^{(1-\alpha)/2\alpha} \rho \sigma^{(1-\alpha)/2\alpha})^\alpha)$ for any density operator ρ , any non-negative operator σ such that $\text{supp}(\rho) \subseteq \text{supp}(\sigma)$ and any $\alpha \in (0, 1) \cup (1, \infty)$. By definition of

⁵More precisely, we require that $\|\text{tr}_G(\rho^i) - \text{tr}_G(\rho^j)\|_1 = o(n^{-2})$.

⁶Note that there exist different quantities that are called topological entanglement entropy (see also [34]).

the measured relative entropy (see (1.6)), we find for any two states ρ and σ

$$\begin{aligned} D_{\mathbb{M}}(\rho\|\sigma) &= \sup_{\mathcal{M} \in \mathbb{M}} D(\mathcal{M}(\rho)\|\mathcal{M}(\sigma)) \geq \sup_{\mathcal{M} \in \mathbb{M}} D_{1/2}(\mathcal{M}(\rho)\|\mathcal{M}(\sigma)) \\ &= -2 \log_2 \inf_{\mathcal{M} \in \mathbb{M}} F(\mathcal{M}(\rho), \mathcal{M}(\sigma)) = -2 \log_2 F(\rho, \sigma), \end{aligned} \quad (4.1)$$

where $\mathbb{M} := \{\mathcal{M} : \mathcal{M}(\rho) = \sum_x \text{tr}(\rho M_x) |x\rangle\langle x| \text{ with } \sum_x M_x = \text{id}\}$ and $\{|x\rangle\}$ is a family of orthonormal vectors. The inequality step uses that $\alpha \mapsto D_{\alpha}(\rho\|\sigma)$ is a monotonically non-decreasing function in α [36, theorem 7] and the final step follows from the fact that for any two states there exists an optimal measurement that does not increase their fidelity [38, §3.3]. As a result, in order to prove theorem 2.1 for finite-dimensional B and C it suffices to prove (2.2).

We first derive a proposition (proposition 4.1) and then show how it can be used to prove (2.2) (and, hence, theorem 2.1). The proposition refers to a family of functions

$$D(A \otimes B \otimes C) \ni \rho \mapsto \Delta_{\mathcal{R}}(\rho) \in \mathbb{R} \cup \{-\infty\}, \quad (4.2)$$

parametrized by recovery maps $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$, where $\text{TPCP}(B, B \otimes C)$ denotes the set of trace-preserving completely positive maps from B to $B \otimes C$ and $D(A \otimes B \otimes C)$ denotes the set of density operators on $A \otimes B \otimes C$. Subsequently in the proof, the function family $\Delta_{\mathcal{R}}(\cdot)$ will be constructed as the difference of the two terms in (2.2) (see equation (4.38)) such that $\Delta_{\mathcal{R}}(\rho) \geq 0$ corresponds to (2.2). The proposition asserts that if for any extension ρ_{ABC} of ρ_{BC} we have $\Delta_{\mathcal{R}}(\rho) \geq 0$ for some $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$ and provided the function family $\Delta_{\mathcal{R}}(\cdot)$ satisfies certain properties described below, then there exists a single recovery map \mathcal{R} for which $\Delta_{\mathcal{R}}(\rho) \geq 0$ for all extensions ρ_{ABC} of ρ_{BC} on a fixed A system. We note that the precise form of the function family $\Delta_{\mathcal{R}}(\cdot)$ is irrelevant for proposition 4.1 as long as it satisfies a list of properties as stated below.

As described above, our goal is to prove that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that $\Delta_{\mathcal{R}}(\rho) \geq 0$ for all $\rho_{ABC} \in D(A \otimes B \otimes C)$ with a fixed marginal ρ_{BC} on $B \otimes C$. To formulate our argument more concisely, we introduce some notation. For any set S of density operators $\rho_{ABC} \in D(A \otimes B \otimes C)$, we define

$$\Delta_{\mathcal{R}}(S) := \inf_{\rho \in S} \Delta_{\mathcal{R}}(\rho). \quad (4.3)$$

The desired statement then reads as $\Delta_{\mathcal{R}}(S) \geq 0$, for any set S of states on $A \otimes B \otimes C$ with a fixed marginal ρ_{BC} . Furthermore, for any fixed states ρ_{ABC}^0 and ρ_{ABC} on $A \otimes B \otimes C$ and $p \in [0, 1]$, we define

$$\rho_{\hat{A}ABC}^p := (1-p)|0\rangle\langle 0|_{\hat{A}} \otimes \rho_{ABC}^0 + p|1\rangle\langle 1|_{\hat{A}} \otimes \rho_{ABC}, \quad (4.4)$$

where \hat{A} is an additional system with two orthogonal states $|0\rangle$ and $|1\rangle$. More generally, for any fixed state ρ_{ABC}^0 and for any set S of density operators ρ_{ABC} we set

$$\mathcal{S}^p := \left\{ \rho_{\hat{A}ABC}^p : \rho_{ABC} \in S \right\}. \quad (4.5)$$

Required properties of the Δ -function 1.

- (i) For any $\rho_{ABC}^0, \rho_{ABC} \in D(A \otimes B \otimes C)$ with identical marginals $\rho_{BC}^0 = \rho_{BC}$ on $B \otimes C$, for any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$, and for any $p \in [0, 1]$ we have $\Delta_{\mathcal{R}}(\rho^p) = (1-p)\Delta_{\mathcal{R}}(\rho^0) + p\Delta_{\mathcal{R}}(\rho)$.
- (ii) For any $\mathcal{R}, \mathcal{R}' \in \text{TPCP}(B, B \otimes C)$, for any $\alpha \in [0, 1]$ and $\bar{\mathcal{R}} = \alpha\mathcal{R} + (1-\alpha)\mathcal{R}'$ we have $\Delta_{\bar{\mathcal{R}}}(\rho) \geq \alpha\Delta_{\mathcal{R}}(\rho) + (1-\alpha)\Delta_{\mathcal{R}'}(\rho)$ for all $\rho \in D(A \otimes B \otimes C)$.
- (iii) For any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$, the function $D(A \otimes B \otimes C) \ni \rho \mapsto \Delta_{\mathcal{R}}(\rho) \in \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous.
- (iv) For any $\rho \in D(A \otimes B \otimes C)$, the function $\text{TPCP}(B, B \otimes C) \ni \mathcal{R} \mapsto \Delta_{\mathcal{R}}(\rho) \in \mathbb{R} \cup \{-\infty\}$ is upper semicontinuous.

Property (i) implies that for any state ρ_{ABC}^0 , for any set S of operators ρ_{ABC} with $\rho_{BC} = \rho_{BC}^0$, and for any $p \in [0, 1]$ we have

$$\Delta_{\mathcal{R}}(S^p) = \inf_{\rho \in S} \Delta_{\mathcal{R}}(\rho^p) = (1-p)\Delta_{\mathcal{R}}(\rho^0) + p \inf_{\rho \in S} \Delta_{\mathcal{R}}(\rho) = (1-p)\Delta_{\mathcal{R}}(\rho^0) + p\Delta_{\mathcal{R}}(S). \quad (4.6)$$

Similarly, property (ii) implies

$$\begin{aligned} \Delta_{\bar{\mathcal{R}}}(S) &= \inf_{\rho \in S} \Delta_{\bar{\mathcal{R}}}(\rho) \geq \inf_{\rho \in S} \{\alpha \Delta_{\mathcal{R}}(\rho) + (1-\alpha)\Delta_{\mathcal{R}'}(\rho)\} \\ &\geq \alpha \inf_{\rho \in S} \Delta_{\mathcal{R}}(\rho) + (1-\alpha) \inf_{\rho \in S} \Delta_{\mathcal{R}'}(\rho) = \alpha \Delta_{\mathcal{R}}(S) + (1-\alpha)\Delta_{\mathcal{R}'}(S). \end{aligned} \quad (4.7)$$

Proposition 4.1. *Let A, B and C be finite-dimensional Hilbert spaces, $\mathcal{P} \subseteq \text{TPCP}(B, B \otimes C)$ be compact and convex, S be a set of density operators on $A \otimes B \otimes C$ with identical marginals on $B \otimes C$, and $\Delta_{\mathcal{R}}(\cdot)$ be a family of functions of the form (4.2) that satisfies properties (i)–(iv). Then*

$$\forall \rho \in S \exists \mathcal{R} \in \mathcal{P} : \Delta_{\mathcal{R}}(\rho) \geq 0 \implies \exists \bar{\mathcal{R}} \in \mathcal{P} : \Delta_{\bar{\mathcal{R}}}(S) \geq 0. \quad (4.8)$$

We now proceed in four steps. In the first, we prove proposition 4.1 for finite sets S . This is done by induction over the cardinality of the set S . We show that if the statement of proposition 4.1 is true for all sets S with $|S| = n$, this implies that it remains valid for all sets S with $|S| = n + 1$. In Step 2, we use an approximation step to extend this to infinite sets S which then completes the proof of proposition 4.1. In the final two steps, we show how to conclude the statement of theorem 2.1 for the finite-dimensional case from that. In Step 3, we prove (2.2) for the case where the recovery map that satisfies (2.2) could still depend on the dimension of the system A . In Step 4, we show how this dependency can be removed.

Proposition 4.1 resembles Sion's minimax theorem [39]. After the completion of this work, it has been noticed that the argument done by proposition 4.1 in this work can be alternatively carried out using Sion's minimax theorem (see [40] for a detailed explanation).

(a) Step 1. Proof of proposition 4.1 for finite size sets S

We proceed by induction over the cardinality $n := |S|$ of the set S of density operators. More precisely, the induction hypothesis is that for any finite-dimensional Hilbert space A and any set S of size n consisting of density operators on $A \otimes B \otimes C$ with fixed marginal ρ_{BC} on $B \otimes C$, the statement (4.8) holds. For $n = 1$, this hypothesis holds trivially for $\bar{\mathcal{R}} = \mathcal{R}$.⁷

We now prove the induction step. Suppose that the induction hypothesis holds for some n . Let A be a finite-dimensional Hilbert space and let $S \cup \{\rho_{ABC}^0\}$ be a set of cardinality $n + 1$ where S is a set of states on $A \otimes B \otimes C$ with fixed marginal ρ_{BC} on $B \otimes C$ of cardinality n and ρ_{ABC}^0 is another state with $\rho_{BC}^0 = \rho_{BC}$. We need to prove that there exists a recovery map $\bar{\mathcal{R}}_{B \rightarrow BC} \in \mathcal{P}$ such that

$$\Delta_{\bar{\mathcal{R}}}(S \cup \{\rho_{ABC}^0\}) \geq 0. \quad (4.9)$$

Let $p \in [0, 1]$ and consider the set S^p as defined in (4.5). In the following, we view the states ρ^p (see equation (4.4)) in this set as tripartite states on $(\hat{A} \otimes A) \otimes B \otimes C$, i.e. we regard the system $\hat{A} \otimes A$ as one (larger) system. The induction hypothesis applied to the extension space $\hat{A} \otimes A$ and the set S^p (of size n) of states on $(\hat{A} \otimes A) \otimes B \otimes C$ implies the existence of a map $\mathcal{R}_{B \rightarrow BC}^p \in \mathcal{P}$ such that

$$\Delta_{\mathcal{R}^p}(S^p) \geq 0. \quad (4.10)$$

As by assumption the function $D(A \otimes B \otimes C) \ni \rho \mapsto \Delta_{\mathcal{R}^p}(\rho) \in \mathbb{R} \cup \{-\infty\}$ satisfies property (i) (and hence also equation (4.6)), we obtain

$$(1-p)\Delta_{\mathcal{R}^p}(\rho^0) + p\Delta_{\mathcal{R}^p}(S) \geq 0. \quad (4.11)$$

⁷For $n = 0$, we have $\Delta_{\mathcal{R}}(S) = \infty \geq 0$ for any $\mathcal{R} \in \mathcal{P}$ since the infimum of an empty set is infinity.

This implies that

$$\Delta_{\mathcal{R}^p}(\rho^0) \geq 0 \quad \text{or} \quad \Delta_{\mathcal{R}^p}(S) \geq 0. \quad (4.12)$$

Furthermore, for $p=0$ the left inequality holds and for $p=1$ the right inequality holds. By choosing $K_0 = \{p \in [0, 1] : \Delta_{\mathcal{R}^p}(\rho^0) \geq 0\}$ and $K_1 = \{p \in [0, 1] : \Delta_{\mathcal{R}^p}(S) \geq 0\}$, the *touching sets lemma* (see lemma 11.1 in the electronic supplementary material) implies that for any $\delta > 0$ there exist $u, v \in [0, 1]$ with $0 \leq v - u \leq \delta$ such that

$$\Delta_{\mathcal{R}^u}(\rho^0) \geq 0 \quad \text{and} \quad \Delta_{\mathcal{R}^v}(S) \geq 0. \quad (4.13)$$

Note also that $\mathcal{R}_{B \rightarrow BC}^u, \mathcal{R}_{B \rightarrow BC}^v \in \mathcal{P}$, as by the induction hypothesis $\mathcal{R}_{B \rightarrow BC}^p \in \mathcal{P}$ for any $p \in [0, 1]$.

We will use this to prove that the recovery map $\tilde{\mathcal{R}} \in \mathcal{P}$ defined by

$$\tilde{\mathcal{R}} := \alpha \mathcal{R}^u + (1 - \alpha) \mathcal{R}^v, \quad (4.14)$$

for an appropriately chosen $\alpha \in [0, 1]$, satisfies

$$\Delta_{\tilde{\mathcal{R}}}(\rho^0) \geq -c\delta \quad \text{and} \quad \Delta_{\tilde{\mathcal{R}}}(S) \geq -c\delta, \quad (4.15)$$

where c is a constant defined by

$$c := 4 \max_{\mathcal{R} \in \text{TPCP}(B, B \otimes C)} \max_{\rho \in D(A \otimes B \otimes C)} \Delta_{\mathcal{R}}(\rho) < \infty. \quad (4.16)$$

Properties (iii) and (iv) together with simple topological facts about the set of density operators and the set of trace-preserving completely positive maps (see lemma 10.1 and remark 10.3 stated in the electronic supplementary material) ensure that the two maxima in (4.16) are attained which implies by the definition of the codomain of $\Delta_{\mathcal{R}}(\cdot)$ (see equation (4.2)) that c is finite. In other words, for any $\delta > 0$ there exists a recovery map $\tilde{\mathcal{R}}^\delta \in \mathcal{P}$ such that

$$\Delta_{\tilde{\mathcal{R}}^\delta}(S \cup \{\rho^0\}) \geq -c\delta. \quad (4.17)$$

The compactness of \mathcal{P} ensures that there exists a recovery map $\tilde{\mathcal{R}} \in \mathcal{P}$ and a sequence $\{\delta_n\}_{n \in \mathbb{N}}$ such that

$$\lim_{n \rightarrow \infty} \delta_n = 0 \quad \text{and} \quad \lim_{n \rightarrow \infty} \tilde{\mathcal{R}}^{\delta_n} = \tilde{\mathcal{R}}. \quad (4.18)$$

Because of (4.17) we have

$$\limsup_{n \rightarrow \infty} \Delta_{\tilde{\mathcal{R}}^{\delta_n}}(S \cup \{\rho^0\}) \geq \lim_{n \rightarrow \infty} -c\delta_n = 0, \quad (4.19)$$

which together with property (iv) implies that

$$\begin{aligned} \Delta_{\tilde{\mathcal{R}}}(S \cup \{\rho^0\}) &= \min_{\rho \in S \cup \{\rho^0\}} \Delta_{\tilde{\mathcal{R}}}(\rho) \geq \min_{\rho \in S \cup \{\rho^0\}} \limsup_{n \rightarrow \infty} \Delta_{\tilde{\mathcal{R}}^{\delta_n}}(\rho) \geq \limsup_{n \rightarrow \infty} \min_{\rho \in S \cup \{\rho^0\}} \Delta_{\tilde{\mathcal{R}}^{\delta_n}}(\rho) \\ &= \limsup_{n \rightarrow \infty} \Delta_{\tilde{\mathcal{R}}^{\delta_n}}(S \cup \{\rho^0\}) \geq 0, \end{aligned} \quad (4.20)$$

and thus proves (4.9).

It thus remains to show (4.15). To simplify the notation, let us define

$$\Lambda^0 := \Delta_{\mathcal{R}^u}(\rho^0) \quad \text{and} \quad \Lambda^1 := \Delta_{\mathcal{R}^v}(S) \quad (4.21)$$

as well as

$$\bar{\Lambda}^0 := \Delta_{\mathcal{R}^v}(\rho^0) \quad \text{and} \quad \bar{\Lambda}^1 := \Delta_{\mathcal{R}^u}(S). \quad (4.22)$$

It follows from (4.11) that

$$(1 - u)\Lambda^0 + u\bar{\Lambda}^1 \geq 0. \quad (4.23)$$

Similarly, we have

$$(1 - v)\bar{\Lambda}^0 + v\Lambda^1 \geq 0. \quad (4.24)$$

As by assumption the function $\Delta_{\mathcal{R}}(\cdot)$ satisfies property (ii) we find together with (4.24) that for any $\alpha \in [0, 1]$ and $\tilde{\mathcal{R}} = \alpha\mathcal{R}^u + (1 - \alpha)\mathcal{R}^v$,

$$\Delta_{\tilde{\mathcal{R}}}(\rho^0) \geq \alpha \Delta_{\mathcal{R}^u}(\rho^0) + (1 - \alpha) \Delta_{\mathcal{R}^v}(\rho^0) = \alpha \Lambda^0 + (1 - \alpha) \bar{\Lambda}^0 \geq \alpha \Lambda^0 - (1 - \alpha) \frac{v}{1 - v} \Lambda^1. \quad (4.25)$$

(If $v = 1$ it suffices to consider the case $\alpha = 1$ so that the last term can be omitted; cf. equation (4.29).) Analogously, using (4.7) and (4.23), we find

$$\Delta_{\tilde{\mathcal{R}}}(\mathcal{S}) \geq \alpha \Delta_{\mathcal{R}^u}(\mathcal{S}) + (1 - \alpha) \Delta_{\mathcal{R}^v}(\mathcal{S}) = \alpha \bar{\Lambda}^1 + (1 - \alpha) \Lambda^1 \geq -\alpha \frac{1 - u}{u} \Lambda^0 + (1 - \alpha) \Lambda^1. \quad (4.26)$$

(If $u = 0$ it suffices to consider the case $\alpha = 0$; cf. equation (4.32).)

To conclude the proof of (4.15), it suffices to choose $\alpha \in [0, 1]$ such that the terms on the right-hand side of (4.25) and (4.26) satisfy

$$\alpha \Lambda^0 - (1 - \alpha) \frac{v}{1 - v} \Lambda^1 \geq -c\delta \quad (4.27)$$

and

$$-\alpha \frac{1 - u}{u} \Lambda^0 + (1 - \alpha) \Lambda^1 \geq -c\delta. \quad (4.28)$$

Let us first assume that $u \geq \frac{1}{2}$. Since Λ^0 and Λ^1 are non-negative (see equation (4.13)), we may choose $\alpha \in [0, 1]$ such that

$$\alpha(1 - v)\Lambda^0 = (1 - \alpha)v\Lambda^1. \quad (4.29)$$

This immediately implies that the left-hand side of (4.27) equals 0, so that the inequality holds. As $\frac{1}{2} \leq u \leq v \leq 1$ and $v - u \leq \delta$ we have

$$\left| \frac{1 - u}{u} - \frac{1 - v}{v} \right| \leq 4\delta. \quad (4.30)$$

Combining this with (4.29), we find

$$-\alpha \frac{1 - u}{u} \Lambda^0 + (1 - \alpha) \Lambda^1 \geq -\alpha \Lambda^0 \left(\frac{1 - v}{v} + 4\delta \right) + (1 - \alpha) \Lambda^1 = -4\alpha \Lambda^0 \delta \geq -4\Lambda^0 \delta, \quad (4.31)$$

which proves (4.28) because by (4.16) we have $\Lambda^0 \leq c/4$.

Analogously, if $u < \frac{1}{2}$, choose $\alpha \in [0, 1]$ such that

$$\alpha(1 - u)\Lambda^0 = (1 - \alpha)u\Lambda^1. \quad (4.32)$$

This immediately implies that the left-hand side of (4.28) equals 0, so that the inequality holds. Furthermore, for $\delta > 0$ sufficiently small such that $v \leq \frac{1}{2}$, we obtain

$$\left| \frac{v}{1 - v} - \frac{u}{1 - u} \right| < 4\delta. \quad (4.33)$$

Together with (4.32) this implies

$$\alpha \Lambda^0 - (1 - \alpha) \frac{v}{1 - v} \Lambda^1 \geq \alpha \Lambda^0 - (1 - \alpha) \Lambda^1 \left(\frac{u}{1 - u} + 4\delta \right) = -4(1 - \alpha) \Lambda^1 \delta \geq -4\Lambda^1 \delta, \quad (4.34)$$

which establishes (4.27). This concludes the proof of proposition 4.1 for sets \mathcal{S} of finite size.

(b) Step 2. Extension to infinite sets \mathcal{S}

All that remains to be done to prove proposition 4.1 is to generalize the statement to arbitrarily large sets \mathcal{S} . In fact, we show that there exists a recovery map $\mathcal{R}_{B \rightarrow BC} \in \mathcal{P}$ such that $\Delta_{\mathcal{R}}(\mathcal{S}) \geq 0$, where \mathcal{S} is the set of all density operators on $A \otimes B \otimes C$ for a fixed finite-dimensional Hilbert space A and a fixed marginal ρ_{BC} .

Note first that this set \mathcal{S} of all density operators on $A \otimes B \otimes C$ with fixed marginal ρ_{BC} on $B \otimes C$ is compact (see lemma 10.2 in the electronic supplementary material). This implies that for any $\varepsilon > 0$ there exists a finite set \mathcal{S}^ε of density operators on $A \otimes B \otimes C$ such that any $\rho \in \mathcal{S}$ is ε -close

to an element of S^ε . We further assume without loss of generality that $S^{\varepsilon'} \subset S^\varepsilon$ for $\varepsilon' \geq \varepsilon$. Let $\mathcal{R}^\varepsilon \in \text{TPCP}(B, B \otimes C)$ be a map such that $\Delta_{\mathcal{R}^\varepsilon}(S^\varepsilon) \geq 0$, whose existence follows from the validity of proposition 4.1 for sets of finite size (which we proved in Step 1). Since the set $\text{TPCP}(B, B \otimes C)$ is compact (see remark 10.3 in the electronic supplementary material) there exists a decreasing sequence $\{\varepsilon_n\}_{n \in \mathbb{N}}$ and $\bar{\mathcal{R}} \in \text{TPCP}(B, B \otimes C)$ such that

$$\lim_{n \rightarrow \infty} \varepsilon_n = 0 \quad \text{and} \quad \bar{\mathcal{R}} = \lim_{n \rightarrow \infty} \mathcal{R}^{\varepsilon_n}. \quad (4.35)$$

Combining this with property (iv) gives for all $n \in \mathbb{N}$

$$\begin{aligned} \Delta_{\bar{\mathcal{R}}}(S^{\varepsilon_n}) &= \inf_{\rho \in S^{\varepsilon_n}} \Delta_{\bar{\mathcal{R}}}(\rho) \geq \inf_{\rho \in S^{\varepsilon_n}} \limsup_{m \rightarrow \infty} \Delta_{\bar{\mathcal{R}}^{\varepsilon_m}}(\rho) \geq \limsup_{m \rightarrow \infty} \inf_{\rho \in S^{\varepsilon_n}} \Delta_{\bar{\mathcal{R}}^{\varepsilon_m}}(\rho) \\ &\geq \limsup_{m \rightarrow \infty} \inf_{\rho \in S^{\varepsilon_m}} \Delta_{\bar{\mathcal{R}}^{\varepsilon_m}}(\rho) = \limsup_{m \rightarrow \infty} \Delta_{\bar{\mathcal{R}}^{\varepsilon_m}}(S^{\varepsilon_m}) \geq 0, \end{aligned} \quad (4.36)$$

where the third inequality holds since $S^{\varepsilon_n} \subset S^{\varepsilon_m}$ for $\varepsilon_n \geq \varepsilon_m$, respectively $n \leq m$. The final inequality follows from the defining property of \mathcal{R}^ε . For any fixed $\rho \in S$ and for all $n \in \mathbb{N}$, let $\rho^n \in S^{\varepsilon_n}$ be such that $\lim_{n \rightarrow \infty} \rho^n = \rho \in S$. (By definition of S^{ε_n} it follows that such a sequence $\{\rho^n\}_{n \in \mathbb{N}}$ with $\rho^n \in S^{\varepsilon_n}$ always exists.) Property (iii) together with (4.36) yields

$$\Delta_{\bar{\mathcal{R}}}(\rho) = \Delta_{\bar{\mathcal{R}}}\left(\lim_{n \rightarrow \infty} \rho^n\right) \geq \limsup_{n \rightarrow \infty} \Delta_{\bar{\mathcal{R}}}(\rho^n) \geq \limsup_{n \rightarrow \infty} \Delta_{\bar{\mathcal{R}}}(S^{\varepsilon_n}) \geq 0. \quad (4.37)$$

Since (4.37) holds for any $\rho \in S$, we obtain $\Delta_{\bar{\mathcal{R}}}(S) \geq 0$, which completes the proof of proposition 4.1.

(c) Step 3. From proposition 4.1 to theorem 2.1 for fixed system A

We next show that theorem 2.1, for the case where A is a fixed finite-dimensional system, follows from proposition 4.1. For this we use proposition 4.1 for the function family

$$\begin{aligned} \Delta_{\mathcal{R}} : D(A \otimes B \otimes C) &\rightarrow \mathbb{R} \cup \{-\infty\} \\ \rho_{ABC} &\mapsto I(A : C|B)_\rho - D_{\mathbb{M}}(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})), \end{aligned} \quad (4.38)$$

with $\mathcal{R}_{B \rightarrow BC} \in \text{TPCP}(B, B \otimes C)$. We note that since C is finite-dimensional this implies that $\Delta_{\mathcal{R}}(\rho) < \infty$ for all $\rho \in D(A \otimes B \otimes C)$. To apply proposition 4.1, we have to verify that the function family $D(A \otimes B \otimes C) \ni \rho \mapsto \Delta_{\mathcal{R}}(\rho) \in \mathbb{R} \cup \{-\infty\}$ of the form (4.38) satisfies the assumptions of the proposition. This is ensured by the following lemma.

Lemma 4.2. *Let A be a separable and B and C finite-dimensional Hilbert spaces. The function family $\Delta_{\mathcal{R}}(\cdot)$ defined by (4.38) satisfies properties (i)–(iv).*

Proof. We first verify that the function $\Delta_{\mathcal{R}}(\cdot)$ satisfies property (i). For any state ρ^p of the form (4.4), we have by the definition of the mutual information

$$I(\hat{A}A : C|B)_{\rho^p} = H(C|B)_{\rho^p} - H(C|BA\hat{A})_{\rho^p}. \quad (4.39)$$

Because $\rho_{BC}^0 = \rho_{BC}$, the first term, $H(C|B)_{\rho^p}$, is independent of p , i.e. $H(C|B)_{\rho^p} = H(C|B)_{\rho^0} = H(C|B)_\rho$. The second term can be written as an expectation over \hat{A} , i.e.

$$H(C|BA\hat{A})_{\rho^p} = (1-p)H(C|BA)_{\rho^0} + pH(C|BA)_{\rho}. \quad (4.40)$$

As a result, we find

$$I(\hat{A}A : C|B)_{\rho^p} = (1-p)I(A : C|B)_{\rho^0} + pI(A : C|B)_\rho. \quad (4.41)$$

The density operator $\mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p)$ can be written as

$$\mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p) = (1-p)|0\rangle\langle 0|_{\hat{A}} \otimes \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0) + p|1\rangle\langle 1|_{\hat{A}} \otimes \mathcal{R}_{B \rightarrow BC}(\rho_{AB}). \quad (4.42)$$

We can thus apply lemma 9.3 given in the electronic supplementary material (which states a linearity property of the measured relative entropy for orthogonal states), from which we obtain

$$D_{\mathbb{M}}\left(\rho_{\hat{A}ABC}^p \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p)\right) = (1-p)D_{\mathbb{M}}\left(\rho_{\hat{A}ABC}^0 \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0)\right) + pD_{\mathbb{M}}\left(\rho_{\hat{A}ABC}^p \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^p)\right). \quad (4.43)$$

Equations (4.41) and (4.43) imply that

$$\Delta_{\mathcal{R}}(\rho^p) = (1-p)\Delta_{\mathcal{R}}(\rho^0) + p\Delta_{\mathcal{R}}(\rho), \quad (4.44)$$

which concludes the proof of property (i).

That $\Delta_{\mathcal{R}}(\cdot)$ satisfies property (ii) can be seen as follows. Let $\mathcal{R}_{B \rightarrow BC}, \mathcal{R}'_{B \rightarrow BC} \in \text{TPCP}(B, B \otimes C)$, $\alpha \in [0, 1]$ and $\tilde{\mathcal{R}}_{B \rightarrow BC} = \alpha\mathcal{R}_{B \rightarrow BC} + (1-\alpha)\mathcal{R}'_{B \rightarrow BC}$. Since the measured relative entropy is convex in the second argument (see lemma 9.4 given in the electronic supplementary material) we find that for any state ρ_{ABC} on $A \otimes B \otimes C$

$$\begin{aligned} D_{\mathbb{M}}(\rho_{ABC} \parallel \tilde{\mathcal{R}}_{B \rightarrow BC}(\rho_{AB})) &= D_{\mathbb{M}}(\rho_{ABC} \parallel \alpha\mathcal{R}_{B \rightarrow BC}(\rho_{AB}) + (1-\alpha)\mathcal{R}'_{B \rightarrow BC}(\rho_{AB})) \\ &\leq \alpha D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + (1-\alpha)D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}'_{B \rightarrow BC}(\rho_{AB})) \end{aligned} \quad (4.45)$$

and hence

$$\Delta_{\tilde{\mathcal{R}}}(\rho) \geq \alpha\Delta_{\mathcal{R}}(\rho) + (1-\alpha)\Delta_{\mathcal{R}'}(\rho). \quad (4.46)$$

We next verify that the function $\Delta_{\mathcal{R}}(\cdot)$ satisfies property (iii). The Alicki–Fannes inequality ensures that $D(A \otimes B \otimes C) \ni \rho \mapsto I(A:C|B)_{\rho} \in \mathbb{R}^+$ is continuous since C is finite-dimensional [35]. By the definition of $\Delta_{\mathcal{R}}(\cdot)$ it thus suffices to show that $D(A \otimes B \otimes C) \ni \rho_{ABC} \mapsto D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) \in \mathbb{R}^+$ is lower semicontinuous. Let $\{\rho_{ABC}^n\}_{n \in \mathbb{N}}$ be a sequence of states on $A \otimes B \otimes C$ such that $\lim_{n \rightarrow \infty} \rho_{ABC}^n = \rho_{ABC} \in D(A \otimes B \otimes C)$. By definition of the measured relative entropy (see (1.6)), we find for $\mathbb{M} := \{\mathcal{M} : \mathcal{M}(\rho) = \sum_x \text{tr}(\rho M_x)|x\rangle\langle x| \text{ with } \sum_x M_x = \text{id}\}$,

$$\begin{aligned} \liminf_{n \rightarrow \infty} D_{\mathbb{M}}(\rho_{ABC}^n \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^n)) &= \liminf_{n \rightarrow \infty} \sup_{\mathcal{M} \in \mathbb{M}} D(\mathcal{M}(\rho_{ABC}^n) \parallel \mathcal{M}(\mathcal{R}_{B \rightarrow BC}(\rho_{AB}^n))) \\ &\geq \sup_{\mathcal{M} \in \mathbb{M}} \liminf_{n \rightarrow \infty} D(\mathcal{M}(\rho_{ABC}^n) \parallel \mathcal{M}(\mathcal{R}_{B \rightarrow BC}(\rho_{AB}^n))) \\ &\geq \sup_{\mathcal{M} \in \mathbb{M}} D(\mathcal{M}(\rho_{ABC}) \parallel \mathcal{M}(\mathcal{R}_{B \rightarrow BC}(\rho_{AB}))) \\ &= D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}_{B \rightarrow BC}(\rho_{AB})). \end{aligned} \quad (4.47)$$

In the penultimate step, we use that the relative entropy is lower semicontinuous [41, Exercise 7.22] and that \mathcal{M} as well as $\mathcal{R}_{B \rightarrow BC}$ are linear and bounded operators and hence continuous.

We finally show that $\Delta_{\mathcal{R}}(\cdot)$ fulfils property (iv). It suffices to verify that $\text{TPCP}(B, B \otimes C) \ni \mathcal{R} \mapsto D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}(\rho_{AB})) \in \mathbb{R}^+$ is lower semicontinuous where by definition of the measured relative entropy (see (1.6)) we have that

$$D_{\mathbb{M}}(\rho_{ABC} \parallel \mathcal{R}(\rho_{AB})) = \sup_{\mathcal{M} \in \mathbb{M}} D(\mathcal{M}(\rho_{ABC}) \parallel \mathcal{M}(\mathcal{R}_{B \rightarrow BC}(\rho_{AB}))). \quad (4.48)$$

Note that since \mathcal{R} and \mathcal{M} are linear bounded operators and hence continuous and the relative entropy for two states σ_1 and σ_2 is defined by $D(\sigma_1 \parallel \sigma_2) := \text{tr}(\sigma_1(\log \sigma_1 - \log \sigma_2))$ we find that $\mathcal{R} \mapsto D(\mathcal{M}(\rho_{ABC}) \parallel \mathcal{M}(\mathcal{R}_{B \rightarrow BC}(\rho_{AB})))$ is continuous as the logarithm $\mathbb{R}^+ \ni x \mapsto \log x \in \mathbb{R}$ is continuous. Since the supremum of continuous functions is lower semicontinuous [42, ch. IV, Section 6.2, Theorem 4], the assertion follows. ■

What remains to be shown in order to apply proposition 4.1 is that for any $\rho \in \mathcal{S}$ where \mathcal{S} is the set of states on $A \otimes B \otimes C$ with a fixed marginal ρ_{BC} on $B \otimes C$, there exists a recovery

map $\mathcal{R}_{B \rightarrow BC} \in \mathcal{P}$ such that $\Delta_{\mathcal{R}}(\rho) \geq 0$. By choosing $\mathcal{P} = \text{TPCP}(B, B \otimes C)$, the main result of [16] however precisely proves this. We have thus shown that $\Delta_{\mathcal{R}}(\rho) \geq 0$ holds for a universal recovery map $\mathcal{R}_{B \rightarrow BC} \in \mathcal{P}$, so that (2.2) follows for any fixed dimension of the A system. This proves the statement of remark 2.2 (and, hence, theorem 2.1) for the case where A is a fixed finite-dimensional Hilbert space.

(d) Step 4. Independence from the A system

Let \mathcal{S} be the set of all density operators on $\bar{A} \otimes B \otimes C$ with a fixed marginal ρ_{BC} on $B \otimes C$, where B and C are finite-dimensional Hilbert spaces and \bar{A} is the infinite-dimensional Hilbert space ℓ^2 of square summable sequences. We now show that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that $\Delta_{\mathcal{R}}(\mathcal{S}) \geq 0$.

Let $\{\Pi_A^a\}_{a \in \mathbb{N}}$ be a sequence of finite-rank projectors on \bar{A} that converges to $\text{id}_{\bar{A}}$ with respect to the weak operator topology. Let \mathcal{S}^a denote the set of states whose marginal on \bar{A} is contained in the support of Π_A^a and with the same fixed marginal ρ_{BC} on $B \otimes C$ as the elements of \mathcal{S} . For all $a \in \mathbb{N}$, let $\mathcal{R}_{B \rightarrow BC}^a$ denote a recovery map that satisfies $\Delta_{\mathcal{R}^a}(\mathcal{S}^a) \geq 0$. Note that the existence of such maps is already established by the proof of theorem 2.1 for the finite-dimensional case. As the set of trace-preserving completely positive maps on finite-dimensional systems is compact (see remark 10.3 in the electronic supplementary material) there exists a subsequence $\{a_i\}_{i \in \mathbb{N}}$ such that $\lim_{i \rightarrow \infty} a_i = \infty$ and $\lim_{i \rightarrow \infty} \mathcal{R}^{a_i} = \bar{\mathcal{R}} \in \text{TPCP}(B, B \otimes C)$. For every $\rho \in \mathcal{S}$, there exists a sequence of states $\{\rho^a\}_{a \in \mathbb{N}}$ with $\rho^a \in \mathcal{S}^a$ that converges to ρ in the trace norm (see lemma 12.3 in the electronic supplementary material). Lemma 4.2 (in particular properties (iii) and (iv)), yields for any $\rho \in \mathcal{S}$

$$\begin{aligned} \Delta_{\bar{\mathcal{R}}}(\rho) &\geq \limsup_{a \rightarrow \infty} \Delta_{\bar{\mathcal{R}}}(\rho^a) \geq \limsup_{a \rightarrow \infty} \limsup_{i \rightarrow \infty} \Delta_{\mathcal{R}^{a_i}}(\rho^a) \geq \limsup_{a \rightarrow \infty} \limsup_{i \rightarrow \infty} \inf_{\rho \in \mathcal{S}^a} \Delta_{\mathcal{R}^{a_i}}(\rho) \\ &\geq \limsup_{i \rightarrow \infty} \inf_{\rho \in \mathcal{S}^{a_i}} \Delta_{\mathcal{R}^{a_i}}(\rho) = \limsup_{i \rightarrow \infty} \Delta_{\mathcal{R}^{a_i}}(\mathcal{S}^{a_i}) \geq 0. \end{aligned} \quad (4.49)$$

The fourth inequality follows since $a_i \geq a$ for large enough i and since this implies that $\mathcal{S}^{a_i} \supset \mathcal{S}^a$, and the final inequality follows by definition of \mathcal{R}^{a_i} . This shows that $\Delta_{\bar{\mathcal{R}}}(\mathcal{S}) \geq 0$.

To retrieve the statement of remark 2.2 (and hence theorem 2.1 for finite-dimensional B and C), we need to argue that this same map $\bar{\mathcal{R}}$ remains valid when we consider any separable space A . In order to do this, observe that any separable Hilbert space A can be isometrically embedded into \bar{A} [21, Theorem II.7]. To conclude, it suffices to remark that $\Delta_{\bar{\mathcal{R}}}$ is invariant under isometries applied on the space A .

5. Extension to infinite dimensions

In this section, we show how to obtain the statement of theorem 2.1 for separable (not necessarily finite-dimensional) Hilbert spaces A, B, C from the finite-dimensional case that has been proven in §4. For trace non-increasing completely positive maps $\mathcal{R}_{B \rightarrow BC}$, we define the function family

$$\begin{aligned} \bar{\Delta}_{\mathcal{R}} : D(A \otimes B \otimes C) &\rightarrow \mathbb{R} \cup \{-\infty\} \\ \rho_{ABC} &\mapsto F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) - 2^{-(1/2)I(A:C|B)_{\rho}}, \end{aligned} \quad (5.1)$$

where $D(A \otimes B \otimes C)$ denotes the set of states on $A \otimes B \otimes C$. We will use the same notation as introduced at the beginning of §4. In addition, we take \mathcal{S} to be the set of all states on $A \otimes B \otimes C$ with a fixed marginal ρ_{BC} on $B \otimes C$. The proof proceeds in two steps where we first show that there exists a sequence of recovery maps $\{\mathcal{R}_{B \rightarrow BC}^k\}_{k \in \mathbb{N}}$ such that $\lim_{k \rightarrow \infty} \bar{\Delta}_{\mathcal{R}^k}(\mathcal{S}) \geq 0$, where the property that all elements of \mathcal{S} have the same marginal on the $B \otimes C$ system will be important. In the second step, we conclude by an approximation argument that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that $\bar{\Delta}_{\mathcal{R}}(\mathcal{S}) \geq 0$.

(a) Step 1. Existence of a sequence of recovery maps

We start by introducing some notation that is used within this step. Let $\{\Pi_B^b\}_{b \in \mathbb{N}}$ and $\{\Pi_C^c\}_{c \in \mathbb{N}}$ be sequences of finite-rank projectors on B and C which converge to id_B and id_C with respect to the weak operator topology. For any given $\rho_{ABC} \in D(A \otimes B \otimes C)$ consider the normalized projected states

$$\rho_{ABC}^{b,c} := \frac{(\text{id}_A \otimes \Pi_B^b \otimes \Pi_C^c) \rho_{ABC} (\text{id}_A \otimes \Pi_B^b \otimes \Pi_C^c)}{\text{tr}((\text{id}_A \otimes \Pi_B^b \otimes \Pi_C^c) \rho_{ABC})} \quad (5.2)$$

and

$$\rho_{ABC}^c := \frac{(\text{id}_A \otimes \text{id}_B \otimes \Pi_C^c) \rho_{ABC} (\text{id}_A \otimes \text{id}_B \otimes \Pi_C^c)}{\text{tr}((\text{id}_A \otimes \text{id}_B \otimes \Pi_C^c) \rho_{ABC})}, \quad (5.3)$$

where for any $c \in \mathbb{N}$, the sequence $\{\rho_{ABC}^{b,c}\}_{b \in \mathbb{N}}$ converges to ρ_{ABC}^c in the trace norm (see, corollary 2 of [43] or lemma 12.1 in the electronic supplementary material) and the sequence $\{\rho_{ABC}^c\}_{c \in \mathbb{N}}$ converges to ρ_{ABC} also in the trace norm. Let $S^{b,c}$ be the set of states that is generated by (5.2) for all $\rho_{ABC} \in \mathcal{S}$. We note that for any given b, c all elements of $S^{b,c}$ have an identical marginal on $B \otimes C$. Let $\mathcal{R}_{B \rightarrow BC}^{b,c}$ denote a recovery map that satisfies $\bar{\Delta}_{\mathcal{R}^{b,c}}(S^{b,c}) \geq 0$ whose existence is established in the proof of theorem 2.1 for finite-dimensional systems B and C (see §4). We next state a lemma that explains how $\bar{\Delta}_{\mathcal{R}}(\rho)$ changes when we replace ρ by a projected state $\rho^{b,c}$.

Lemma 5.1. *For any $\rho_{BC} \in D(B \otimes C)$ there exists a sequence of reals $\{\xi^{b,c}\}_{b,c \in \mathbb{N}}$ with⁸ $\lim_{c \rightarrow \infty} \lim_{b \rightarrow \infty} \xi^{b,c} = 0$, such that for any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$, any extension ρ_{ABC} of ρ_{BC} , and $\rho_{ABC}^{b,c}$ as given in (5.2) we have*

$$\bar{\Delta}_{\mathcal{R}}(\rho^{b,c}) - \bar{\Delta}_{\mathcal{R}}(\rho) \leq \xi^{b,c} \quad \text{for all } b, c \in \mathbb{N}. \quad (5.4)$$

Proof. We note that local projections applied to the subsystem C can only decrease the mutual information, i.e.

$$\text{tr}(\Pi_C^c \rho_C) I(A : C|B)_{\rho^c} \leq I(A : C|B)_{\rho}. \quad (5.5)$$

To see this assume that a measurement with respect to Π_C^c as well as its orthogonal complement is applied to ρ . Furthermore, let Z be a random variable that stores the outcome of this measurement. Then by the data processing inequality

$$\begin{aligned} I(A : C|B)_{\rho} &= H(A|B)_{\rho} - H(A|CB)_{\rho} \geq H(A|B)_{\rho'} - H(A|CBZ)_{\rho'} \\ &\geq H(A|BZ)_{\rho'} - H(A|CBZ)_{\rho'} = I(A : C|BZ)_{\rho'}, \end{aligned} \quad (5.6)$$

where ρ' is the state after the measurement. Because $I(A : C|BZ)_{\rho'}$ can be written as the expectation over the mutual information of the post-measurement states conditioned on the different values of Z , and because all these terms are non-negative, the above claim follows.

The Alicki–Fannes inequality [35] ensures that for a fixed finite-dimensional system C the conditional mutual information $I(A : C|B)_{\rho} = H(C|B)_{\rho} - H(C|AB)_{\rho}$ is continuous in ρ with respect to the trace norm, i.e.

$$I(A : C|B)_{\rho^{b,c}} \leq I(A : C|B)_{\rho^c} + 8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c}), \quad (5.7)$$

where $\varepsilon^{b,c} = \|\rho_{ABC}^{b,c} - \rho_{ABC}^c\|_1$ and $h(\cdot)$ denotes the binary Shannon entropy function defined by $h(p) := -p \log_2(p) - (1-p) \log_2(1-p)$ for $0 \leq p \leq 1$. Using the Fuchs–van de Graaf inequality [44]

⁸The precise form of the sequence $\{\xi^{b,c}\}_{b,c \in \mathbb{N}}$ is given in the proof (see equation (5.17)).

and a variant of the gentle-measurement lemma (see lemma 12.1 given in the electronic supplementary material), we find

$$\varepsilon^{b,c} \leq 2\sqrt{1 - F\left(\rho_{ABC}^{b,c}, \rho_{ABC}^c\right)^2} \leq 2\sqrt{1 - \frac{\text{tr}\left(\Pi_B^b \otimes \Pi_C^c \rho_{BC}\right)}{\text{tr}(\Pi_C^c \rho_C)}}. \quad (5.8)$$

Combining (5.5) and (5.7) yields

$$I(A : C|B)_{\rho^{b,c}} \leq \frac{1}{\text{tr}(\Pi_C^c \rho_C)} I(A : C|B)_\rho + 8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c}). \quad (5.9)$$

Since $x^y \leq x - y + 1$ for $x, y \in [0, 1]$,⁹ we find

$$\begin{aligned} 2^{-(1/2)I(A:C|B)_\rho} &= 2^{-(1/2)I(A:C|B)_{\rho^{b,c}}} \\ &\leq 2^{-(1/2)I(A:C|B)_\rho} - 2^{-(1/2)\text{tr}(\Pi_C^c \rho_C)I(A:C|B)_{\rho^{b,c}} - \text{tr}(\Pi_C^c \rho_C)} + 1. \end{aligned} \quad (5.10)$$

According to (5.9) and since $2^{-x} \geq 1 - \ln(2)x$ for $x \in \mathbb{R}$, we have

$$\begin{aligned} 2^{-(1/2)\text{tr}(\Pi_C^c \rho_C)I(A:C|B)_{\rho^{b,c}}} &\geq 2^{-(1/2)I(A:C|B)_\rho} 2^{-(1/2)\text{tr}(\Pi_C^c \rho_C)(8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c}))} \\ &\geq 2^{-(1/2)I(A:C|B)_\rho} - \frac{\ln(2)}{2} \text{tr}(\Pi_C^c \rho_C) \left(8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c})\right). \end{aligned} \quad (5.11)$$

Combining (5.10) and (5.11) yields

$$\begin{aligned} 2^{-(1/2)I(A:C|B)_\rho} &= 2^{-(1/2)I(A:C|B)_{\rho^{b,c}}} \\ &\leq \frac{\ln(2)}{2} \text{tr}(\Pi_C^c \rho_C) \left(8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c})\right) + (1 - \text{tr}(\Pi_C^c \rho_C)) \\ &\leq \frac{\ln(2)}{2} \left(8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c})\right) + (1 - \text{tr}(\Pi_C^c \rho_C)). \end{aligned} \quad (5.12)$$

For two states σ_1 and σ_2 let $P(\sigma_1, \sigma_2) := \sqrt{1 - F(\sigma_1, \sigma_2)^2}$ denote the purified distance. Applying the Fuchs–van de Graaf inequality [44] and a variant of the gentle-measurement lemma (see lemma 12.1 in the electronic supplementary material) gives

$$P\left(\rho_{ABC}, \rho_{ABC}^{b,c}\right)^2 = 1 - F\left(\rho_{ABC}, \rho_{ABC}^{b,c}\right)^2 \leq 1 - \text{tr}\left(\Pi_B^b \otimes \Pi_C^c \rho_{BC}\right). \quad (5.13)$$

Since the purified distance is a metric [45] that is monotonous under trace-preserving completely positive maps [46, theorem 3.4], (5.13) gives

$$\begin{aligned} P(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) &\leq P\left(\rho_{ABC}, \rho_{ABC}^{b,c}\right) + P\left(\rho_{ABC}^{b,c}, \mathcal{R}_{B \rightarrow BC}\left(\rho_{AB}^{b,c}\right)\right) \\ &\quad + P\left(\mathcal{R}_{B \rightarrow BC}\left(\rho_{AB}^{b,c}\right), \mathcal{R}_{B \rightarrow BC}(\rho_{AB})\right) \\ &\leq 2P\left(\rho_{ABC}, \rho_{ABC}^{b,c}\right) + P\left(\rho_{ABC}^{b,c}, \mathcal{R}_{B \rightarrow BC}\left(\rho_{AB}^{b,c}\right)\right) \\ &\leq P\left(\rho_{ABC}^{b,c}, \mathcal{R}_{B \rightarrow BC}\left(\rho_{AB}^{b,c}\right)\right) + 2\sqrt{1 - \text{tr}\left(\Pi_B^b \otimes \Pi_C^c \rho_{BC}\right)}. \end{aligned} \quad (5.14)$$

⁹For $x = 0$ the statement clearly holds. For $(0, 1] \times [0, 1] \ni (x, y) \mapsto f(x, y) := x^y - x + y - 1 \in \mathbb{R}$ we find by using the convexity of $y \mapsto f(x, y)$ that $\max_{x \in (0, 1]} \max_{y \in [0, 1]} f(x, y) = 0$.

As the fidelity for states lies between zero and one, (5.14) implies

$$\begin{aligned}
 & F(\rho_{ABC}^{b,c}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^{b,c}))^2 \\
 & \leq F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))^2 + 4 \left(1 - \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC})\right) + 4 \sqrt{1 - \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC})} \\
 & \leq F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))^2 + 8 \sqrt{1 - \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC})} \\
 & \leq \left(F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + 2\sqrt{2} \left(1 - \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC})\right)^{1/4}\right)^2.
 \end{aligned} \tag{5.15}$$

This implies that

$$F(\rho_{ABC}^{b,c}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^{b,c})) \leq F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + 2\sqrt{2}(1 - \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC}))^{1/4}. \tag{5.16}$$

By definition of the quantity $\bar{\Delta}_{\mathcal{R}}(\cdot)$ (see equation (5.1)) the combination of (5.12) and (5.16) yields

$$\begin{aligned}
 & \bar{\Delta}_{\mathcal{R}}(\rho^{b,c}) - \bar{\Delta}_{\mathcal{R}}(\rho) \\
 & \leq \frac{\ln(2)}{2} (8\varepsilon^{b,c} \log(\text{rank } \Pi_C^c) + 4h(\varepsilon^{b,c})) + (1 - \text{tr}(\Pi_C^c \rho_C)) + 2\sqrt{2}(1 - \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC}))^{1/4} \\
 & =: \xi^{b,c},
 \end{aligned} \tag{5.17}$$

where $\varepsilon^{b,c}$ is bounded by (5.8). By a variant of the gentle-measurement lemma (see lemma 12.2 in the electronic supplementary material), we find $\lim_{b \rightarrow \infty} \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC}) = \text{tr}(\Pi_C^c \rho_C)$ for all $c \in \mathbb{N}$ and hence $\lim_{b \rightarrow \infty} \varepsilon^{b,c} = 0$ for any $c \in \mathbb{N}$. Furthermore, we have $\lim_{c \rightarrow \infty} \text{tr}(\Pi_C^c \rho_C) = 1$ and $\lim_{c \rightarrow \infty} \lim_{b \rightarrow \infty} \text{tr}(\Pi_B^b \otimes \Pi_C^c \rho_{BC}) = 1$ which implies that $\lim_{c \rightarrow \infty} \lim_{b \rightarrow \infty} \xi^{b,c} = 0$. This proves the assertion. ■

By lemma 5.1, using the notation defined at the beginning of Step 1, we find

$$\begin{aligned}
 \limsup_{c \rightarrow \infty} \limsup_{b \rightarrow \infty} \bar{\Delta}_{\mathcal{R}^{b,c}}(S) &= \limsup_{c \rightarrow \infty} \limsup_{b \rightarrow \infty} \inf_{\rho \in S} \bar{\Delta}_{\mathcal{R}^{b,c}}(\rho) \\
 &\geq \limsup_{c \rightarrow \infty} \limsup_{b \rightarrow \infty} \inf_{\rho \in S} \{\bar{\Delta}_{\mathcal{R}^{b,c}}(\rho^{b,c}) - \xi^{b,c}\} \\
 &= \limsup_{c \rightarrow \infty} \limsup_{b \rightarrow \infty} \left\{ \inf_{\rho^{b,c} \in S^{b,c}} \bar{\Delta}_{\mathcal{R}^{b,c}}(\rho^{b,c}) \right\} - \xi^{b,c} \\
 &= \limsup_{c \rightarrow \infty} \limsup_{b \rightarrow \infty} \bar{\Delta}_{\mathcal{R}^{b,c}}(S^{b,c}) \\
 &\geq 0,
 \end{aligned} \tag{5.18}$$

where the second equality step is valid since all states in S have the same fixed marginal on $B \otimes C$ and since the sequence $\{\xi^{b,c}\}_{b,c \in \mathbb{N}}$ only depends on this marginal. The penultimate step uses that $\lim_{c \rightarrow \infty} \lim_{b \rightarrow \infty} \xi^{b,c} = 0$. The final inequality follows by definition of $\mathcal{R}_{B \rightarrow BC}^{b,c}$. Inequality (5.18) implies that there exist sequences $\{b_k\}_{k \in \mathbb{N}}$ and $\{c_k\}_{k \in \mathbb{N}}$ such that $\limsup_{k \rightarrow \infty} \Delta_{\mathcal{R}^{b_k, c_k}}(S) \geq 0$. Setting $\mathcal{R}_{B \rightarrow BC}^k = \mathcal{R}_{B \rightarrow BC}^{b_k, c_k}$ then implies that there exists a sequence $\{\mathcal{R}_{B \rightarrow BC}^k\}_{k \in \mathbb{N}}$ of recovery maps that satisfies

$$\limsup_{k \rightarrow \infty} \bar{\Delta}_{\mathcal{R}^k}(S) \geq 0. \tag{5.19}$$

(b) Step 2. Existence of a limit

Recall that \mathcal{S} is the set of density operators on $A \otimes B \otimes C$ with a fixed marginal ρ_{BC} on $B \otimes C$. The goal of this step is to use (5.19) to prove that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that

$$\bar{\Delta}_{\mathcal{R}}(\mathcal{S}) \geq 0. \quad (5.20)$$

Let $\{\Pi_B^m\}_{m \in \mathbb{N}}$ and $\{\Pi_C^m\}_{m \in \mathbb{N}}$ be sequences of projectors with rank m that weakly converge to id_B and id_C , respectively. Furthermore, for any m and any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$ let $[\mathcal{R}]^m$ be the trace non-increasing map obtained from \mathcal{R} by projecting the input and output with Π_B^m and $\Pi_B^m \otimes \Pi_C^m$, respectively. We start with a preparatory lemma that proves a relationship between $\bar{\Delta}_{[\mathcal{R}]^m}(\mathcal{S})$ and $\bar{\Delta}_{\mathcal{R}}(\mathcal{S})$.

Lemma 5.2. *For any $\rho_{BC} \in D(B \otimes C)$ there exists a sequence of reals $\{\delta^m\}_{m \in \mathbb{N}}$ with $\lim_{m \rightarrow \infty} \delta^m = 0$,¹⁰ such that for any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$ we have*

$$\bar{\Delta}_{[\mathcal{R}]^m}(\mathcal{S}) \geq \bar{\Delta}_{\mathcal{R}}(\mathcal{S}) - \delta^m - 4\varepsilon^{1/4}, \quad (5.21)$$

where $\|\mathcal{R}(\rho_B) - \rho_{BC}\|_1 \leq \varepsilon$.

Proof. For any $\rho_{ABC} \in \mathcal{S}$ and any $m \in \mathbb{N}$ let us define the non-negative operator $\hat{\rho}_{AB}^m := (\text{id}_A \otimes \Pi_B^m) \rho_{AB} (\text{id}_A \otimes \Pi_B^m)$. By definition of $\bar{\Delta}_{\mathcal{R}}(\cdot)$ (see equation (5.1)), it suffices to show that for any $\rho_{ABC} \in \mathcal{S}$, any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$, $\varepsilon \in [0, 2]$ such that $\|\mathcal{R}(\rho_B) - \rho_{BC}\|_1 \leq \varepsilon$ and

$$\hat{\rho}_{ABC}^m := (\text{id}_A \otimes \Pi_B^m \otimes \Pi_C^m) \mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m) (\text{id}_A \otimes \Pi_B^m \otimes \Pi_C^m) \quad (5.22)$$

we have $F(\rho_{ABC}, \hat{\rho}_{ABC}^m) \geq F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) - \delta^m - 4\varepsilon^{1/4}$. As in Step 1, let $P(\cdot, \cdot)$ denote the purified distance. A variant of the gentle-measurement lemma (see lemma 12.1 in the electronic supplementary material) implies that

$$P(\rho_{AB}, \hat{\rho}_{AB}^m)^2 = 1 - F(\rho_{AB}, \hat{\rho}_{AB}^m)^2 \leq 1 - \text{tr}(\rho_B \Pi_B^m)^2. \quad (5.23)$$

Similarly, we obtain

$$\begin{aligned} P(\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m), \hat{\rho}_{ABC}^m)^2 &\leq 1 - \text{tr}(\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m) \Pi_B^m \otimes \Pi_C^m)^2 \\ &= 1 - \text{tr}(\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_B^m) \Pi_B^m \otimes \Pi_C^m)^2. \end{aligned} \quad (5.24)$$

By Hölder's inequality, monotonicity of the trace norm for trace-preserving completely positive maps [47, example 9.1.8 and corollary 9.1.10] and (5.23) together with the Fuchs–van de Graaf inequality [44] and a variant of the gentle-measurement lemma (see lemma 12.1 given in the electronic supplementary material), we find

$$\begin{aligned} |\text{tr}((\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_B^m) - \mathcal{R}_{B \rightarrow BC}(\rho_B)) \Pi_B^m \otimes \Pi_C^m)| &\leq \|\mathcal{R}_{B \rightarrow BC}(\rho_B) - \mathcal{R}_{B \rightarrow BC}(\hat{\rho}_B^m)\|_1 \|\Pi_B^m \otimes \Pi_C^m\|_\infty \\ &= \|\mathcal{R}_{B \rightarrow BC}(\rho_B) - \mathcal{R}_{B \rightarrow BC}(\hat{\rho}_B^m)\|_1 \leq \|\rho_B - \hat{\rho}_B^m\|_1 \\ &\leq \|\rho_{AB} - \hat{\rho}_{AB}^m\|_1 \leq 2\sqrt{1 - \text{tr}(\rho_B \Pi_B^m)^2}. \end{aligned} \quad (5.25)$$

Combining (5.24), (5.25) and Hölder's inequality together with the assumption $\|\mathcal{R}(\rho_B) - \rho_{BC}\|_1 \leq \varepsilon$ gives

$$\begin{aligned} P(\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m), \hat{\rho}_{ABC}^m)^2 &\leq 1 - \text{tr}(\mathcal{R}_{B \rightarrow BC}(\rho_B) \Pi_B^m \otimes \Pi_C^m)^2 + 4\sqrt{1 - \text{tr}(\rho_B \Pi_B^m)^2} \\ &\leq 1 - \text{tr}(\rho_{BC} \Pi_B^m \otimes \Pi_C^m)^2 + 4\sqrt{1 - \text{tr}(\rho_B \Pi_B^m)^2} + 2\varepsilon. \end{aligned} \quad (5.26)$$

¹⁰The precise form of the sequence $\{\delta^m\}_{m \in \mathbb{N}}$ can be found in the proof (see equation (5.28)).

Inequalities (5.23), (5.26) and the monotonicity of the purified distance under trace-preserving and completely positive maps [46, Theorem 3.4] show that

$$\begin{aligned} P(\rho_{ABC}, \tilde{\rho}_{ABC}^m) &\leq P(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + P(\mathcal{R}_{B \rightarrow BC}(\rho_{AB}), \mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m)) + P(\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m), \tilde{\rho}_{ABC}^m) \\ &\leq P(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + P(\rho_{AB}, \hat{\rho}_{AB}^m) + P(\mathcal{R}_{B \rightarrow BC}(\hat{\rho}_{AB}^m), \tilde{\rho}_{ABC}^m) \\ &\leq P(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + \frac{(\delta^m)^2}{8} + \sqrt{2\varepsilon}, \end{aligned} \quad (5.27)$$

for

$$\delta^m := \sqrt{8} \left(\sqrt{1 - \text{tr}(\rho_B \Pi_B^m)^2} + \sqrt{1 - \text{tr}(\rho_{BC} \Pi_B^m \otimes \Pi_C^m)^2} + 4\sqrt{1 - \text{tr}(\rho_B \Pi_B^m)^2} \right)^{1/2}. \quad (5.28)$$

As the purified distance between two states lies inside the interval $[0, 1]$ and since $(\delta^m)^2/8 + \sqrt{2\varepsilon} \in [0, 6]$, (5.27) implies that whenever $F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))^2 \geq (\delta^m)^2 + 8\sqrt{2\varepsilon}$, we have

$$\begin{aligned} F(\rho_{ABC}, \tilde{\rho}_{ABC}^m)^2 &\geq F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))^2 - (\delta^m)^2 - 8\sqrt{2\varepsilon} \\ &\geq (F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) - \sqrt{(\delta^m)^2 + 8\sqrt{2\varepsilon}})^2. \end{aligned} \quad (5.29)$$

As a result, we find

$$F(\rho_{ABC}, \tilde{\rho}_{ABC}^m) \geq F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) - \delta^m - \sqrt{8(2\varepsilon)^{1/4}}, \quad (5.30)$$

which proves (5.21) since $\sqrt{8}2^{1/4} \leq 4$.

Recall that B and C are separable Hilbert spaces and that $\{\Pi_B^m\}_{m \in \mathbb{N}}$ and $\{\Pi_B^m \otimes \Pi_C^m\}_{m \in \mathbb{N}}$ converge weakly to id_B and $\text{id}_B \otimes \text{id}_C$, respectively. A variant of the gentle-measurement lemma (see lemma 12.2 given in the electronic supplementary material) thus shows that $\lim_{m \rightarrow \infty} \delta^m = 0$ since $\lim_{m \rightarrow \infty} \text{tr}(\rho_B \Pi_B^m) = 1$ and $\lim_{m \rightarrow \infty} \text{tr}(\rho_{BC} \Pi_B^m \otimes \Pi_C^m) = 1$. ■

The following lemma proves that for sufficiently large m and a recovery map $\mathcal{R}_{B \rightarrow BC}$ that maps ρ_B to density operators that are close to ρ_{BC} , the operator $[\mathcal{R}]^m(\rho_{AB})$ has a trace that is bounded from below by essentially one.

Lemma 5.3. *Let A, B and C be separable Hilbert spaces. For any density operator $\rho_{AB} \in D(A \otimes B)$ and any $\mathcal{R} \in \text{TPCP}(B, B \otimes C)$, we have*

$$\text{tr}([\mathcal{R}]^m(\rho_{AB})) \geq \text{tr}(\Pi_B^m \otimes \Pi_C^m \rho_{BC}) - 2\sqrt{1 - \text{tr}(\Pi_B^m \rho_B) - \|\mathcal{R}(\rho_B) - \rho_{BC}\|_1}. \quad (5.31)$$

Proof. We first note that by Hölder's inequality and monotonicity of the trace norm for trace-preserving completely positive maps [47, example 9.1.8 and corollary 9.1.10] we have

$$\begin{aligned} &|\text{tr}(\Pi_B^m \otimes \Pi_C^m (\mathcal{R}(\rho_B) - \mathcal{R}(\Pi_B^m \rho_B \Pi_B^m)))| \\ &\leq \|\mathcal{R}(\rho_B) - \mathcal{R}(\Pi_B^m \rho_B \Pi_B^m)\|_1 \leq \|\rho_B - \Pi_B^m \rho_B \Pi_B^m\|_1. \end{aligned} \quad (5.32)$$

Together with Hölder's inequality this implies

$$\begin{aligned} \text{tr}([\mathcal{R}]^m(\rho_{AB})) &= \text{tr}(\Pi_B^m \otimes \Pi_C^m \mathcal{R}(\Pi_B^m \rho_{AB} \Pi_B^m)) = \text{tr}(\Pi_B^m \otimes \Pi_C^m \mathcal{R}(\Pi_B^m \rho_B \Pi_B^m)) \\ &\geq \text{tr}(\Pi_B^m \otimes \Pi_C^m \mathcal{R}(\rho_B)) - \|\rho_B - \Pi_B^m \rho_B \Pi_B^m\|_1 \\ &\geq \text{tr}(\Pi_B^m \otimes \Pi_C^m \rho_{BC}) - \|\rho_B - \Pi_B^m \rho_B \Pi_B^m\|_1 - \|\mathcal{R}(\rho_B) - \rho_{BC}\|_1. \end{aligned} \quad (5.33)$$

Combining a generalization of the Fuchs–van de Graaf inequality (see lemma 8.2 in the electronic supplementary material) and a variant of the gentle-measurement lemma (see lemma 12.1 in the electronic supplementary material) gives

$$\begin{aligned}\|\rho_B - \Pi_B^m \rho_B \Pi_B^m\|_1 &\leq 2\sqrt{1 - F(\rho_B, \Pi_B^m \rho_B \Pi_B^m)^2} = 2\sqrt{1 - \text{tr}(\Pi_B^m \rho_B) F\left(\frac{\rho_B, \Pi_B^m \rho_B \Pi_B^m}{\text{tr}(\Pi_B^m \rho_B)}\right)^2} \\ &\leq 2\sqrt{1 - \text{tr}(\Pi_B^m \rho_B)},\end{aligned}\quad (5.34)$$

which together with (5.33) proves the assertion. \blacksquare

According to (5.19) the mappings \mathcal{R}^k satisfy

$$\bar{\Delta}_{\mathcal{R}^k}(S) \geq -\tilde{\varepsilon}^k, \quad (5.35)$$

with $\tilde{\varepsilon}^k \geq 0$ such that $\liminf_{k \rightarrow \infty} \tilde{\varepsilon}^k = 0$. As explained in remark 2.3, by considering a state $\bar{\rho}_{ABC} = \rho_A \otimes \rho_B \in \mathcal{S}$, (5.35) implies $F(\rho_{BC}, \mathcal{R}^k(\rho_B)) \geq -\tilde{\varepsilon}^k + 1$. Applying the Fuchs–van de Graaf inequality [44] gives

$$\|\rho_{BC} - \mathcal{R}^k(\rho_B)\|_1 \leq 2\sqrt{\tilde{\varepsilon}^k(2 - \tilde{\varepsilon}^k)} =: \varepsilon^k, \quad (5.36)$$

where $\liminf_{k \rightarrow \infty} \varepsilon^k = 0$ because $\liminf_{k \rightarrow \infty} \tilde{\varepsilon}^k = 0$.

By lemma 5.2, we have

$$\bar{\Delta}_{[\mathcal{R}^k]^m}(S) \geq \bar{\Delta}_{\mathcal{R}^k}(S) - 4(\varepsilon^k)^{1/4} - \delta^m. \quad (5.37)$$

Hence, using our starting point (5.19),

$$\limsup_{k \rightarrow \infty} \bar{\Delta}_{[\mathcal{R}^k]^m}(S) \geq \limsup_{k \rightarrow \infty} \bar{\Delta}_{\mathcal{R}^k}(S) - 4(\varepsilon^k)^{1/4} - \delta^m \geq -\delta^m. \quad (5.38)$$

Because, for any fixed $m \in \mathbb{N}$, the mappings $[\mathcal{R}^k]^m$, for $k \in \mathbb{N}$, are all contained in the same finite-dimensional subspace (i.e. the set of trace non-increasing maps from operators on the support of Π_B^m to operators on the support of $\Pi_B^m \otimes \Pi_C^m$), and because the space of all such mappings is compact (see remark 10.3 in the electronic supplementary material), for any fixed $m \in \mathbb{N}$ there exists a subsequence of the sequence $\{[\mathcal{R}^k]^m\}_{k \in \mathbb{N}}$ that converges. Specifically for any fixed $m \in \mathbb{N}$ there exists a sequence $\{k_i^m\}_{i \in \mathbb{N}}$ such that

$$\bar{\mathcal{R}}^m := \lim_{i \rightarrow \infty} [\mathcal{R}^{k_i^m}]^m \quad (5.39)$$

is well defined. Furthermore, because of the continuity of $\mathcal{R} \mapsto \bar{\Delta}_{\mathcal{R}}(\rho_{ABC})$ on the set of maps from operators on the support of Π_B^m to operators on the support of $\Pi_B^m \otimes \Pi_C^m$ (see lemma 10.4 given in the electronic supplementary material), we have

$$\begin{aligned}\bar{\Delta}_{\bar{\mathcal{R}}^m}(S) &= \inf_{\rho \in \mathcal{S}} \bar{\Delta}_{\bar{\mathcal{R}}^m}(\rho) = \inf_{\rho \in \mathcal{S}} \lim_{i \rightarrow \infty} \bar{\Delta}_{[\mathcal{R}^{k_i^m}]^m}(\rho) \geq \limsup_{i \rightarrow \infty} \inf_{\rho \in \mathcal{S}} \bar{\Delta}_{[\mathcal{R}^{k_i^m}]^m}(\rho) \\ &= \limsup_{i \rightarrow \infty} \bar{\Delta}_{[\mathcal{R}^{k_i^m}]^m}(S) \geq -\delta^m,\end{aligned}\quad (5.40)$$

and, hence,

$$\liminf_{m \rightarrow \infty} \bar{\Delta}_{\bar{\mathcal{R}}^m}(S) \geq 0. \quad (5.41)$$

Without loss of generality, we can assume that the projector Π_B^m is in the eigenbasis of ρ_B that is denoted by $\{|b\rangle_B^m\}$. For a basis $\{|b\rangle_B^m\}$, we define the projector $\Pi_B^m = W \Pi_B^m W^\dagger$ for an isometry $W = \sum_b |b\rangle |b\rangle_B^m$. For any $m \in \mathbb{N}$, let $\rho_{BC;\bar{B}}^m$ be the operator obtained by applying $\bar{\mathcal{R}}^m$ to a purification $\rho_{B;\bar{B}} = (\rho_B^{1/2} \otimes \text{id}_{\bar{B}}) \sum_b |b\rangle_B \otimes |b\rangle_{\bar{B}}$ of ρ_B .

As explained above, there exists a converging subsequence $\{k_i^{m+1}\}_{i \in \mathbb{N}}$ of $\{k_i^m\}_{i \in \mathbb{N}}$ such that $\tilde{\mathcal{R}}^m := \lim_{i \rightarrow \infty} [\mathcal{R}^{k_i^{m+1}}]^m$. Using the definition of $\tilde{\mathcal{R}}^m$ and that $\Pi_B^m \leq \Pi_B^{m'}$, $\Pi_C^m \leq \Pi_C^{m'}$ and $\Pi_{\bar{B}}^m \leq \Pi_{\bar{B}}^{m'}$ for $m \leq m'$, we obtain

$$\begin{aligned} \rho_{BC;\bar{B}}^m &= \tilde{\mathcal{R}}^m(\rho_{B;\bar{B}}) \\ &= \lim_{i \rightarrow \infty} [\mathcal{R}^{k_i^{m+1}}]^m(\rho_{B;\bar{B}}) = \lim_{i \rightarrow \infty} (\Pi_B^m \otimes \Pi_C^m) [\mathcal{R}^{k_i^{m+1}}]^{m+1} (\Pi_B^m \rho_{B;\bar{B}} \Pi_B^m) (\Pi_B^m \otimes \Pi_C^m) \\ &= \lim_{i \rightarrow \infty} (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) [\mathcal{R}^{k_i^{m+1}}]^{m+1} (\rho_{B;\bar{B}}) (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) \\ &= (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) \tilde{\mathcal{R}}^{m+1}(\rho_{B;\bar{B}}) (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) \\ &= (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) \rho_{BC;\bar{B}}^{m+1} (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m). \end{aligned} \quad (5.42)$$

As a result, since $\Pi_B^m \leq \Pi_B^{m'}$, $\Pi_C^m \leq \Pi_C^{m'}$ and $\Pi_{\bar{B}}^m \leq \Pi_{\bar{B}}^{m'}$ for $m \leq m'$, we have for any $m \leq m'$

$$\rho_{BC;\bar{B}}^m = (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) \rho_{BC;\bar{B}}^{m'} (\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m). \quad (5.43)$$

A variant of the gentle-measurement lemma (see lemma 12.1 in the electronic supplementary material) together with (5.43) implies

$$\begin{aligned} F(\rho_{BC;\bar{B}}^m, \rho_{BC;\bar{B}}^{m'}) &= F(\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m \rho_{BC;\bar{B}}^{m'} \Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m, \rho_{BC;\bar{B}}^m) \\ &\geq \text{tr}(\rho_{BC;\bar{B}}^{m'} \Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m) = \text{tr}(\rho_{BC;\bar{B}}^m). \end{aligned} \quad (5.44)$$

A generalization of the Fuchs–van de Graaf inequality (see lemma 8.2 in the electronic supplementary material) yields for $m' \geq m$

$$\|\rho_{BC;\bar{B}}^m - \rho_{BC;\bar{B}}^{m'}\|_1 \leq 2\sqrt{\text{tr}(\rho_{BC;\bar{B}}^{m'})^2 - F(\rho_{BC;\bar{B}}^m, \rho_{BC;\bar{B}}^{m'})^2} \leq 2\sqrt{\text{tr}(\rho_{BC;\bar{B}}^{m'})^2 - \text{tr}(\rho_{BC;\bar{B}}^m)^2}. \quad (5.45)$$

We now prove that as $m \rightarrow \infty$, $\text{tr}(\rho_{BC;\bar{B}}^m)$ goes to 1. Note that since B is a separable Hilbert space and $\rho_{B;\bar{B}}$ is normalized it can be written as $\rho_{B;\bar{B}} = |\psi\rangle\langle\psi|$, where $|\psi\rangle$ is a state on $B \otimes \bar{B}$. Furthermore, as $\Pi_B^m \otimes \Pi_C^m \otimes \Pi_{\bar{B}}^m \leq \text{id}_{BC\bar{B}}$, (5.43) implies that

$$\text{tr}(\rho_{BC;\bar{B}}^m) \leq \text{tr}(\rho_{BC;\bar{B}}^{m'}) \leq 1 \quad \text{for } m' \geq m. \quad (5.46)$$

By definition of $\rho_{BC;\bar{B}}^m$, lemma 5.3 together with (5.36) implies that

$$\begin{aligned} \lim_{m \rightarrow \infty} \text{tr}(\rho_{BC;\bar{B}}^m) &= \lim_{m \rightarrow \infty} \lim_{i \rightarrow \infty} \text{tr}([\mathcal{R}^{k_i^m}]^m(\rho_{B;\bar{B}})) \\ &\geq \lim_{m \rightarrow \infty} \text{tr}(\Pi_B^m \otimes \Pi_C^m \rho_{BC}) - \|\rho_B - \Pi_B^m \rho_B \Pi_B^m\|_1 - \liminf_{i \rightarrow \infty} \varepsilon^{k_i^m} \\ &\geq \lim_{m \rightarrow \infty} \text{tr}(\Pi_B^m \otimes \Pi_C^m \rho_{BC}) - 2\sqrt{1 - \text{tr}(\Pi_B^m \rho_B)^2} = 1, \end{aligned} \quad (5.47)$$

where the second inequality uses a generalized version of the Fuchs–van de Graaf inequality (see lemma 8.2 in the electronic supplementary material), a variant of the gentle-measurement lemma (see lemma 12.1 in the electronic supplementary material), and that $\liminf_{i \rightarrow \infty} \varepsilon^{k_i^m} = 0$ for all $m \in \mathbb{N}$. The final step follows by another variant of the gentle-measurement lemma (see lemma 12.2 in the electronic supplementary material).

Equations (5.45)–(5.47) show that, $\{\rho_{BC;\bar{B}}^m\}_{m \in \mathbb{N}}$ is a Cauchy sequence. Because the set of sub-normalized non-negative operators (i.e. the set of sub-normalized density operators) is complete,¹¹ this sequence converges towards such an operator, i.e. we can define a

¹¹We note that the set of sub-normalized density operators on a Hilbert space is clearly closed. Since every Hilbert space is complete and as every closed subspace of a complete space is complete [42, ch. II, section 3.4, proposition 8] this implies that the set of sub-normalized density operators is complete.

density operator

$$\tilde{\rho}_{BC:\tilde{B}} := \lim_{m \rightarrow \infty} \rho_{BC:\tilde{B}}^m. \quad (5.48)$$

We note that the operators $\rho_{BC:\tilde{B}}^m$ are not normalized in general. However, (5.47) shows that $\tilde{\rho}_{BC:\tilde{B}}$ has unit trace. We now define the recovery map $\mathcal{R}_{B \rightarrow BC}$ as the one that maps $\rho_{B:\tilde{B}}$ to $\tilde{\rho}_{BC:\tilde{B}}$. We note that this does not uniquely define the recovery map $\mathcal{R}_{B \rightarrow BC}$, which is not a problem as theorem 2.1 proves the *existence* of a recovery map that satisfies (2.1) and does not claim that this map is unique. It remains to show that $\mathcal{R}_{B \rightarrow BC}$ has the property (5.20). This follows from the observation that any density operator ρ_{AB} can be obtained from the purification $\rho_{B:\tilde{B}}$ by applying a trace-preserving completely positive map $\mathcal{T}_{\tilde{B} \rightarrow A}$ from \tilde{B} to A . By a continuity property stated in lemma 10.5 in the electronic supplementary material and because $\mathcal{T}_{\tilde{B} \rightarrow A}$ commutes with any recovery map $\mathcal{R}_{B \rightarrow BC}$ from B to $B \otimes C$, we have

$$\begin{aligned} \mathcal{R}_{B \rightarrow BC}(\rho_{AB}) &= (\mathcal{R}_{B \rightarrow BC} \circ \mathcal{T}_{\tilde{B} \rightarrow A})(\rho_{B:\tilde{B}}) = (\mathcal{T}_{\tilde{B} \rightarrow A} \circ \mathcal{R}_{B \rightarrow BC})(\rho_{B:\tilde{B}}) = \mathcal{T}_{\tilde{B} \rightarrow A}(\tilde{\rho}_{BC:\tilde{B}}) \\ &= \mathcal{T}_{\tilde{B} \rightarrow A}(\lim_{m \rightarrow \infty} \rho_{BC:\tilde{B}}^m) = \lim_{m \rightarrow \infty} \mathcal{T}_{\tilde{B} \rightarrow A}(\rho_{BC:\tilde{B}}^m) = \lim_{m \rightarrow \infty} (\mathcal{T}_{\tilde{B} \rightarrow A} \circ \tilde{\mathcal{R}}_{B \rightarrow BC}^m)(\rho_{B:\tilde{B}}) \\ &= \lim_{m \rightarrow \infty} (\tilde{\mathcal{R}}_{B \rightarrow BC}^m \circ \mathcal{T}_{\tilde{B} \rightarrow A})(\rho_{B:\tilde{B}}) = \lim_{m \rightarrow \infty} \tilde{\mathcal{R}}_{B \rightarrow BC}^m(\rho_{AB}). \end{aligned} \quad (5.49)$$

Using the continuity of the fidelity [12,13], this implies that

$$\bar{\Delta}_{\mathcal{R}}(\rho) = \lim_{m \rightarrow \infty} \bar{\Delta}_{\tilde{\mathcal{R}}^m}(\rho), \quad (5.50)$$

for any $\rho \in \mathcal{S}$. Combining this with (5.41) gives

$$\bar{\Delta}_{\mathcal{R}}(\mathcal{S}) = \inf_{\rho \in \mathcal{S}} \bar{\Delta}_{\mathcal{R}}(\rho) = \inf_{\rho \in \mathcal{S}} \lim_{m \rightarrow \infty} \bar{\Delta}_{\tilde{\mathcal{R}}^m}(\rho) \geq \liminf_{m \rightarrow \infty} \inf_{\rho \in \mathcal{S}} \bar{\Delta}_{\tilde{\mathcal{R}}^m}(\rho) = \liminf_{m \rightarrow \infty} \bar{\Delta}_{\tilde{\mathcal{R}}^m}(\mathcal{S}) \geq 0, \quad (5.51)$$

which concludes Step 2 and thus completes the proof of theorem 2.1 in the general case where B and C are no longer finite-dimensional.

6. Proof of corollary 2.4

The first statement of corollary 2.4 that holds for separable Hilbert spaces follows immediately from theorem 2.1, since $2^{-(1/2)I(A:C|B)_\rho} \geq 1 - (\ln(2)/2)I(A:C|B)_\rho$. The proof of the second statement of corollary 2.4 is partitioned into three steps.¹² We first show that a similar method as used in §4 can be used to reveal certain insights about the structure of the recovery map $\mathcal{R}_{B \rightarrow BC}$ (which is not universal) that satisfies

$$F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) \geq 1 - \frac{\ln(2)}{2} I(A:C|B)_\rho. \quad (6.1)$$

In a second step, by invoking proposition 4.1, we use this knowledge to prove that for a fixed A system there exists a recovery map that satisfies (6.1) which is universal and preserves the structure of the non-universal recovery map from before. Finally, in Step 3 we show how the dependency on the fixed A system can be removed.

(a) Step 1. Structure of a non-universal recovery map

We will show that for any density operator ρ_{ABC} on $A \otimes B \otimes C$, where A , B , and C are finite-dimensional Hilbert spaces there exists a trace-preserving completely positive map $\mathcal{R}_{B \rightarrow BC}$ that satisfies (6.1) and is of the form

$$X_B \mapsto \rho_{BC}^{1/2} W_{BC} (\rho_B^{-1/2} X_B \rho_B^{-1/2} \otimes \text{id}_C) W_{BC}^\dagger \rho_{BC}^{1/2}, \quad (6.2)$$

on the support of ρ_B , where W_{BC} is a unitary on $B \otimes C$. We start by proving the following preparatory lemma.

¹²Although corollary 2.4 does not immediately follow from theorem 2.1 it is justified to term it as such, as it follows by the same proof technique that is used to derive theorem 2.1 (in particular it makes use of proposition 4.1).

Lemma 6.1. For any density operator ρ_{ABC} on $A \otimes B \otimes C$, where A , B and C are finite-dimensional Hilbert spaces there exists a trace-preserving completely positive map $\mathcal{R}_{B \rightarrow BC}$ of the form

$$X_B \mapsto V_{BC} \rho_{BC}^{1/2} (\rho_B^{-1/2} U_B X_B U_B^\dagger \rho_B^{-1/2} \otimes \text{id}_C) \rho_{BC}^{1/2} V_{BC}^\dagger, \quad (6.3)$$

where V_{BC} is a unitary on $B \otimes C$ that commutes with ρ_{BC} and U_B is a unitary on B that commutes with ρ_B such that

$$F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) \geq 1 - \frac{\ln(2)}{2} I(A : C|B)_\rho. \quad (6.4)$$

Proof. Let ρ_{ABC} be an arbitrary state on $A \otimes B \otimes C$ and let ρ_{ABC}^0 be a Markov chain with the same marginal on the $B \otimes C$ system, i.e. $\rho_{BC}^0 = \rho_{BC}$. For $p \in (0, 1]$, define the state

$$\rho_{\hat{A}ABC}^p := (1-p)|0\rangle\langle 0|_{\hat{A}} \otimes \rho_{ABC}^0 + p|1\rangle\langle 1|_{\hat{A}} \otimes \rho_{ABC}. \quad (6.5)$$

The main result of [6] (see theorem 5.1 and remark 4.3 in [6]) implies that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ of the form

$$X_B \mapsto V_{BC} \rho_{BC}^{1/2} (\rho_B^{-1/2} U_B X_B U_B^\dagger \rho_B^{-1/2} \otimes \text{id}_C) \rho_{BC}^{1/2} V_{BC}^\dagger, \quad (6.6)$$

where U_B is diagonal with respect to the eigenbasis of ρ_B , $U_B U_B^\dagger \leq \text{id}_B$ and V_{BC} is a unitary on $B \otimes C$, such that

$$F\left(\rho_{\hat{A}ABC}^p, \mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p)\right) \geq 1 - \frac{\ln(2)}{2} I(\hat{A}A : C|B)_{\rho^p}. \quad (6.7)$$

(Alternatively this statement also follows from [20]—which however appeared after the completion of this work.) By lemma 6.2, using that $I(A : C|B)_{\rho^0} = 0$ since ρ_{ABC}^0 is a Markov chain, this may be rewritten as

$$\begin{aligned} & p(1 - F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))) + (1-p)(1 - F(\rho_{ABC}^0, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0))) \\ & \leq p \frac{\ln(2)}{2} I(A : C|B)_\rho. \end{aligned} \quad (6.8)$$

Let us assume by contradiction that any recovery map $\mathcal{R}_{B \rightarrow BC}$ that satisfies (6.8) does not leave ρ_{ABC}^0 invariant, i.e. $\rho_{ABC}^0 \neq \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0)$. This implies that there exists a $\delta_{\mathcal{R}} \in (0, 1]$, which may depend on the recovery map $\mathcal{R}_{B \rightarrow BC}$, such that $1 - F(\rho_{ABC}^0, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0)) = \delta_{\mathcal{R}}$. In the following, we argue that there exists a universal (i.e. independent of $\mathcal{R}_{B \rightarrow BC}$) constant $\delta \in (0, 1]$ such that $1 - F(\rho_{ABC}^0, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0)) \geq \delta$ for all recovery maps $\mathcal{R}_{B \rightarrow BC}$ that satisfy (6.8). Since the set of trace-preserving completely positive maps from B to $B \otimes C$ that satisfy (6.8) is compact¹³ and the function $f : \text{TPCP}(B, B \otimes C) \ni \mathcal{R}_{B \rightarrow BC} \mapsto 1 - F(\rho_{ABC}^0, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0)) \in [0, 1]$ is continuous (see lemma 10.4 in the electronic supplementary material), Weierstrass' theorem ensures that $\delta := \min_{\mathcal{R}_{B \rightarrow BC}} f(\mathcal{R}_{B \rightarrow BC})$, where we optimize over the set of trace-preserving completely positive maps from B to $B \otimes C$ that satisfy (6.8), exists. By assumption, for every recovery map $\mathcal{R}_{B \rightarrow BC}$ that satisfies (6.8) we have $f(\mathcal{R}_{B \rightarrow BC}) > 0$ and hence $\delta \in (0, 1]$. If we insert any such recovery map $\mathcal{R}_{B \rightarrow BC}$ into (6.8), this gives

$$1 - F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + \frac{\delta}{p} - \delta \leq \frac{\ln(2)}{2} I(A : C|B)_\rho, \quad (6.9)$$

which cannot be valid for sufficiently small p . To see this, we note that (6.9) can be rewritten as

$$p \geq \frac{\delta}{(\ln(2)/2)I(A : C|B)_\rho + \delta + F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) - 1}, \quad (6.10)$$

since C is assumed to be a finite-dimensional system and as such $I(A : C|B)_\rho < \infty$. This contradicts our assumption that every recovery map that satisfies (6.8) does not leave ρ_{ABC}^0 invariant. Since by [6] for any $p \in (0, 1]$ there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ of the form (6.3) that satisfies (6.8) we

¹³This set is bounded as the set of trace-preserving completely positive maps from B to $B \otimes C$ is bounded (see remark 10.3 in the electronic supplementary material). Furthermore, this set is closed since the set of trace-preserving completely positive maps from B to $B \otimes C$ is closed (see remark 10.3 in the electronic supplementary material) and the mapping $\mathcal{R}_{B \rightarrow BC} \mapsto F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))$ is continuous for all states ρ_{ABC} (see lemma 10.4 in the electronic supplementary material). The Heine–Borel theorem then implies compactness.

conclude that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ of the form (6.3) that satisfies (6.8) and leaves ρ_{ABC}^0 invariant. We note that for recovery maps that leave ρ_{ABC}^0 invariant, (6.8) simplifies to (6.4) for all p . Thus, there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ of the form (6.3) satisfying (6.4) that leaves ρ_{ABC}^0 invariant, i.e. $\mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0) = \rho_{ABC}^0$. Since $\rho_{ABC}^0 := \rho_A \otimes \rho_{BC}$ is a Markov chain with marginal $\rho_{BC}^0 = \rho_{BC}$, the condition $\mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0) = \rho_{ABC}^0$ implies that $\mathcal{R}_{B \rightarrow BC}(\rho_B) = \rho_{BC}$.

We have thus shown that there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ that satisfies (6.4) and fulfils

$$\mathcal{R}_{B \rightarrow BC}(\rho_B) = V_{BC} \rho_{BC}^{1/2} (U_B U_B^\dagger \otimes \text{id}_C) \rho_{BC}^{1/2} V_{BC}^\dagger = \rho_{BC}. \quad (6.11)$$

Using the fact that $\mathcal{R}_{B \rightarrow BC}$ is trace preserving shows that

$$\text{id}_B = \text{tr}_C(U_B^\dagger \rho_B^{-1/2} \rho_{BC}^{1/2} V_{BC}^\dagger \rho_{BC}^{1/2} \rho_B^{-1/2} U_B) = U_B^\dagger \rho_B^{-1/2} \rho_B \rho_B^{-1/2} U_B = U_B^\dagger U_B. \quad (6.12)$$

This simplifies (6.11) to $V_{BC} \rho_{BC} V_{BC}^\dagger = \rho_{BC}$, i.e. V_{BC} and ρ_{BC} commute which concludes the proof. ■

Lemma 6.1 implies that the mapping (6.3) can be written as

$$X_B \mapsto \rho_{BC}^{1/2} W_{BC} (\rho_B^{-1/2} X_B \rho_B^{-1/2} \otimes \text{id}_C) W_{BC}^\dagger \rho_{BC}^{1/2}, \quad (6.13)$$

with $W_{BC} = V_{BC} U_B \otimes \text{id}_C$ which is a unitary as V_{BC} and U_B are unitaries. Furthermore, W_{BC} is such that (6.13) is trace-preserving.

(b) Step 2. Structure of a universal recovery map for fixed A system

In this step, we show that the recovery map satisfying (6.1) of the form (6.2), whose existence has been established in Step 1, can be made universal without sacrificing the (partial) knowledge about its structure. The idea is to apply proposition 4.1 for the function family

$$\begin{aligned} \tilde{\Delta}_{\mathcal{R}}(\rho) : D(A \otimes B \otimes C) &\rightarrow \mathbb{R} \cup \{-\infty\} \\ \rho_{ABC} &\mapsto F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) - 1 + \frac{\ln(2)}{2} I(A : C|B)_\rho. \end{aligned} \quad (6.14)$$

We therefore need to verify that the assumptions of proposition 4.1 are fulfilled. This is done by the following lemma. We first note that since C is finite-dimensional this implies that $\tilde{\Delta}_{\mathcal{R}}(\rho) < \infty$ for all $\rho \in D(A \otimes B \otimes C)$.

Lemma 6.2. *Let A be a separable and B and C finite-dimensional Hilbert spaces. The function family $\tilde{\Delta}_{\mathcal{R}}(\cdot)$ defined by (6.14) satisfies properties (i)–(iv).*

Proof. We start by showing that $\tilde{\Delta}_{\mathcal{R}}(\cdot)$ satisfies property (i). For $\rho_{\hat{A}ABC}^p$ as defined in (4.4), we claim

$$F(\rho_{\hat{A}ABC}^p, \mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p)) = (1-p)F(\rho_{ABC}^0, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0)) + pF(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})). \quad (6.15)$$

The density operator $\mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p)$ can be written as

$$\mathcal{R}_{B \rightarrow BC}(\rho_{\hat{A}AB}^p) = (1-p)|0\rangle\langle 0|_{\hat{A}} \otimes \mathcal{R}_{B \rightarrow BC}(\rho_{AB}^0) + p|1\rangle\langle 1|_{\hat{A}} \otimes \mathcal{R}_{B \rightarrow BC}(\rho_{AB}). \quad (6.16)$$

The relevant density operators thus satisfy the orthogonality conditions for equality in lemma 8.1 given in the electronic supplementary material, from which (6.15) follows. Furthermore, as explained in the proof of lemma 4.2 we have

$$I(\hat{A}A : C|B)_{\rho^p} = (1-p)I(A : C|B)_{\rho^0} + pI(A : C|B)_\rho. \quad (6.17)$$

Equations (6.15) and (6.17) imply that

$$\tilde{\Delta}_{\mathcal{R}}(\rho^p) = (1-p)\tilde{\Delta}_{\mathcal{R}}(\rho^0) + p\tilde{\Delta}_{\mathcal{R}}(\rho). \quad (6.18)$$

We next verify that $\tilde{\Delta}_{\mathcal{R}}(\cdot)$ fulfils property (ii). Let $\mathcal{R}_{B \rightarrow BC}, \mathcal{R}'_{B \rightarrow BC} \in \text{TPCP}(B, B \otimes C)$, $\alpha \in [0, 1]$ and $\tilde{\mathcal{R}}_{B \rightarrow BC} = \alpha \mathcal{R}_{B \rightarrow BC} + (1 - \alpha) \mathcal{R}'_{B \rightarrow BC}$. A specific property of the fidelity stated in lemma 8.1 in the electronic supplementary material implies that for any state ρ_{ABC} on $A \otimes B \otimes C$

$$\begin{aligned} F(\rho_{ABC}, \tilde{\mathcal{R}}_{B \rightarrow BC}(\rho_{AB})) &= F(\rho_{ABC}, \alpha \mathcal{R}_{B \rightarrow BC}(\rho_{AB}) + (1 - \alpha) \mathcal{R}'_{B \rightarrow BC}(\rho_{AB})) \\ &\geq \alpha F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB})) + (1 - \alpha) F(\rho_{ABC}, \mathcal{R}'_{B \rightarrow BC}(\rho_{AB})), \end{aligned} \quad (6.19)$$

and, hence, by the definition of $\tilde{\Delta}_{\mathcal{R}}(\cdot)$

$$\tilde{\Delta}_{\tilde{\mathcal{R}}}(\rho) \geq \alpha \tilde{\Delta}_{\mathcal{R}}(\rho) + (1 - \alpha) \tilde{\Delta}_{\mathcal{R}'}(\rho). \quad (6.20)$$

The function $\rho \mapsto \tilde{\Delta}_{\mathcal{R}}(\rho)$ is continuous which clearly implies property (iii). To see this, recall that by the Alicki–Fannes inequality $\rho \mapsto I(A : C|B)_{\rho}$ is continuous for a finite-dimensional C system [35]. Furthermore, since $\rho_{AB} \mapsto \mathcal{R}_{BC}(\rho_{AB})$ is continuous (see lemma 10.5 in the electronic supplementary material), the continuity of the fidelity [12,13] implies that $\rho_{ABC} \mapsto F(\rho_{ABC}, \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))$ is continuous, which then establishes property (iii).

Finally, it remains to show that $\tilde{\Delta}_{\mathcal{R}}(\cdot)$ satisfies property (iv), which however follows directly by lemma 10.4 given in the electronic supplementary material. ■

Let $\mathcal{P} \subseteq \text{TPCP}(B, B \otimes C)$ be the convex hull of the set of trace-preserving completely positive mappings from the B to the $B \otimes C$ system that are of the form (6.2). We note that the elements of \mathcal{P} are mappings of the form (2.4), since a convex combination of unitary mappings are unital and a convex combination of trace-preserving maps remains trace-preserving. Proposition 4.1, which is applicable as shown in lemma 6.2 together with Step 1 therefore proves the assertion for a fixed finite-dimensional A system.

(c) Step 3. Independence from the A system

Let S be the set of all density operators on $\tilde{A} \otimes B \otimes C$ with a fixed marginal ρ_{BC} on $B \otimes C$, where B and C are finite-dimensional Hilbert spaces and \tilde{A} is the infinite-dimensional Hilbert space ℓ^2 of square summable sequences.

We note that the set of trace-preserving completely positive maps of the form (2.4) on finite-dimensional systems is compact, which follows by remark 10.3 (see the electronic supplementary material) together with the fact that the intersection of a compact set and a closed set is compact. Hence, using lemma 6.2 (in particular properties (iii) and (iv)) and the result from Step 2, the same argument as in Step 4 of §4 can be applied to conclude the existence of a recovery map $\mathcal{R}_{B \rightarrow BC}$ of the form (2.4) such that $\tilde{\Delta}_{\mathcal{R}}(S) \geq 0$.

As every separable Hilbert space A can isometrically be embedded into \tilde{A} [21, Theorem II.7] and since $\tilde{\Delta}_{\mathcal{R}}$ is invariant under isometries applied on the extension space A , we can conclude that the recovery map $\mathcal{R}_{B \rightarrow BC}$ remains valid for any separable extension space A . This proves the statement of corollary 2.4 for finite-dimensional B and C systems.

7. Discussion

Our main result is that for any density operator ρ_{BC} on $B \otimes C$ there exists a recovery map $\mathcal{R}_{B \rightarrow BC}$ such that the distance between *any* extension ρ_{ABC} of ρ_{BC} acting on $A \otimes B \otimes C$ and $\mathcal{R}_{B \rightarrow BC}(\rho_{AB})$ is bounded from above by the conditional mutual information $I(A : C|B)_{\rho}$. It is natural to ask whether such a map can be described as a simple and explicit function of ρ_{BC} . In fact, it was conjectured in [5,11] that (1.2) holds for a very simple choice of map, namely

$$\mathcal{T}_{B \rightarrow BC} : X_B \mapsto \rho_{BC}^{1/2} (\rho_B^{-1/2} X_B \rho_B^{-1/2} \otimes \text{id}_C) \rho_{BC}^{1/2}, \quad (7.1)$$

called the *transpose map* or *Petz recovery map*. This conjecture, if correct, would have important consequences in obtaining remainder terms for the monotonicity of the relative entropy [19]. As discussed in the Introduction, if ρ_{ABC} is such that it is a (perfect) quantum Markov chain or the B system is classical, the claim of the conjecture is known to hold.

One possible approach to prove a result of this form would be to start from the result (1.2) for an unknown recovery map and then show that the transpose map $\mathcal{T}_{B \rightarrow BC}$ cannot be much worse than any other recovery map. In fact, a theorem of Barnum & Knill [48] directly implies that when ρ_{ABC} is pure, we have

$$F(\rho_{ABC}, \mathcal{T}_{B \rightarrow BC}(\rho_{AB})) \leq F(A; C|B)_\rho \leq \sqrt{F(\rho_{ABC}, \mathcal{T}_{B \rightarrow BC}(\rho_{AB}))}. \quad (7.2)$$

This shows that, if ρ_{ABC} is pure, an inequality of the form (1.2), with the fidelity replaced by its square root, holds for the transpose map. In order to generalize this to all states, one might hope that (7.2) also holds for mixed states ρ_{ABC} . However, this turns out to be wrong even when the state ρ_{ABC} is completely classical (see §13 in the electronic supplementary material for an example).

Another interesting question is whether the lower bound in terms of the measured relative entropy (2.2) can be improved to a relative entropy. Such an inequality is known to be false if we restrict the recovery map to be the transpose map (7.1) [9], but it might be true when we optimize over all recovery maps. It is worth noting that in case such an inequality holds for any ρ_{ABC} and a corresponding recovery map, then the argument presented in this work would imply that there exists a universal recovery map satisfying (2.2) with the relative entropy instead of the measured relative entropy. This can be seen by defining the function family $\rho \mapsto \Delta_{\mathcal{R}}(\rho) := I(A : C|B)_\rho - D(\rho_{ABC} \| \mathcal{R}_{B \rightarrow BC}(\rho_{AB}))$. A linearity property of the relative entropy for orthogonal states (see lemma 9.2 in the electronic supplementary material), the convexity of the relative entropy [49, theorem 11.12] and the lower semicontinuity of the relative entropy [41, example 7.22] imply that $\Delta_{\mathcal{R}}(\cdot)$ satisfies properties (i)–(iv). As a result, proposition 4.1 is applicable which can be used to prove the existence of a universal recovery map.

After the completion of this work, there was a series of works around finding improvements or alternative proofs for inequality (1.2). In [20], an alternative proof for (1.2) based on the Hadamard three-line theorem was discovered.¹⁴ After that yet another proof for (1.2) has been found which is based elementary properties of pinching maps and the operator logarithm [50]. Finally in [40] an *explicit* and universal recovery map has been determined that satisfies (1.2) based on Hirschman's strengthening [51] of the Hadamard three-line theorem.

Data accessibility. This work does not have any experimental data.

Authors' contributions. All authors contributed equally to this work.

Competing interests. We have no competing interests.

Funding. This project was supported by the European Research Council (ERC) via grant no. 258932, by the Swiss National Science Foundation (SNSF) via the National Centre of Competence in Research 'QSIT', and by the European Commission via the project 'RAQUEL'.

Acknowledgements. We thank Mario Berta, Fernando Brandão, Philipp Kammerlander, Joseph Renes, Volker Scholz, Marco Tomamichel and Mark Wilde for discussions about approximate Markov chains.

References

1. Petz D. 1986 Sufficient subalgebras and the relative entropy of states of a von Neumann algebra. *Commun. Math. Phys.* **105**, 123–131. (doi:10.1007/BF01212345)
2. Petz D. 1986 Sufficiency of channels over von Neumann algebras. *Q. J. Math. Oxf.* **39**, 97–108. (doi:10.1093/qmath/39.1.97)
3. Hayden P, Jozsa R, Petz D, Winter A. 2004 Structure of states which satisfy strong subadditivity of quantum entropy with equality. *Commun. Math. Phys.* **246**, 359–374. (doi:10.1007/s00220-004-1049-z)
4. Datta N, Wilde MM. 2015 Quantum Markov chains, sufficiency of quantum channels, and Rényi information measures. *J. Phys. A: Math. Theor.* **48**, 505301. (doi:10.1088/1751-8113/48/50/505301)

¹⁴The result in [20] is more general than (1.2) as it proves a remainder term for the monotonicity of the relative entropy which implies (1.2) as a special case (see [20] for more details).

5. Kim IH. 2013 Application of conditional independence to gapped quantum many-body systems. See <http://www.physics.usyd.edu.au/quantum/Coogee2013/Presentations/Kim.pdf>.
6. Fawzi O, Renner R. 2015 Quantum conditional mutual information and approximate Markov chains. *Commun. Math. Phys.* **340**, 575–611. (doi:10.1007/s00220-015-2466-x)
7. Christandl M, Schuch N, Winter A. 2012 Entanglement of the antisymmetric state. *Commun. Math. Phys.* **311**, 397–422. (doi:10.1007/s00220-012-1446-7)
8. Ibinson B, Linden N, Winter A. 2008 Robustness of quantum Markov chains. *Commun. Math. Phys.* **277**, 289–304. (doi:10.1007/s00220-007-0362-8)
9. Winter A, Li K. 2012 A stronger subadditivity relation? With applications to squashed entanglement, sharability and separability. See http://www.maths.bris.ac.uk/csajw/stronger_subadditivity.pdf.
10. Zhang L. 2013 Conditional mutual information and commutator. *Int. J. Theor. Phys.* **52**, 2112–2117. (doi:10.1007/s10773-013-1505-7)
11. Berta M, Seshadreesan KP, Wilde MM. 2015 Rényi generalizations of the conditional quantum mutual information. *J. Math. Phys.* **56**, 022205. (doi:10.1063/1.4908102)
12. Uhlmann A. 1976 The ‘transition probability’ in the state space of a $*$ -algebra. *Rep. Math. Phys.* **9**, 273–279. (doi:10.1016/0034-4877(76)90060-4)
13. Jozsa R. 1994 Fidelity for mixed quantum states. *J. Mod. Opt.* **41**, 2315–2323. (doi:10.1080/09500349414552171)
14. Seshadreesan KP, Wilde MM. 2015 Fidelity of recovery, squashed entanglement, and measurement recoverability. *Phys. Rev. A* **92**, 042321. (doi:10.1103/PhysRevA.92.042321)
15. Berta M, Tomamichel M. 2015 The fidelity of recovery is multiplicative. (<http://arxiv.org/abs/1502.07973>)
16. Brandão FGSL, Harrow AW, Oppenheim J, Strelchuk S. 2015 Quantum conditional mutual information, reconstructed states, and state redistribution. *Phys. Rev. Lett.* **115**, 050501. (doi:10.1103/PhysRevLett.115.050501)
17. Hiai F, Petz D. 1991 The proper formula for relative entropy and its asymptotics in quantum probability. *Commun. Math. Phys.* **143**, 99–114. (doi:10.1007/BF02100287)
18. Hayashi M. 2001 Asymptotics of quantum relative entropy from a representation theoretical viewpoint. *J. Phys. A Math. Gen.* **34**, 3413. (doi:10.1088/0305-4470/34/16/309)
19. Berta M, Lemm M, Wilde M. 2015 Monotonicity of quantum relative entropy and recoverability. *Quant. Inform. Comput.* **15**, 1333–1354.
20. Wilde MM. 2015 Recoverability in quantum information theory. *Proc. R. Soc. A* **471**, 20150338. (doi:10.1098/rspa.2015.0338)
21. Reed M, Simon B. 1980 *Functional analysis*. Amsterdam, The Netherlands: Elsevier.
22. Einsiedler M, Ward T. 2010 *Ergodic theory*. Berlin, Germany: Springer.
23. Lieb E, Ruskai M. 1973 Proof of the strong subadditivity of quantum-mechanical entropy. *J. Math. Phys.* **14**, 1938–1941. (doi:10.1063/1.1666274)
24. Lieb EH, Ruskai MB. 1973 A fundamental property of quantum-mechanical entropy. *Phys. Rev. Lett.* **30**, 434–436. (doi:10.1103/PhysRevLett.30.434)
25. Li K, Winter A. 2014 Squashed entanglement, k -extendibility, quantum Markov chains, and recovery maps. (<http://arxiv.org/abs/1410.4184>)
26. Kim IH. 2013 Conditional independence in quantum many-body systems. PhD thesis, Caltech.
27. Brandão FG, Harrow AW. 2013 Product-state approximations to quantum ground states. In *Proc. of the 45th Annual ACM Symp. on Theory of Computing. STOC '13, Palo Alto, CA, 1–4 June*, pp. 871–880. New York, NY: ACM.
28. Brandão FG, Harrow AW. 2013 Quantum de Finetti theorems under local measurements with applications. In *Proc. of the 45th Annual ACM Symposium on Theory of Computing. STOC '13, Palo Alto, CA, 1–4 June*, pp. 861–870. New York, NY: ACM.
29. Jain R, Radhakrishnan J, Sen P. 2003 A lower bound for the bounded round quantum communication complexity of set disjointness. In *IEEE 44th Annual IEEE Symp. on Foundations of Computer Science, Cambridge, MA, 11–14 October*, pp. 220–229. Piscataway, NJ: IEEE.
30. Kerenidis I, Laplante S, Lerays V, Roland J, Xiao D. 2012 Lower bounds on information complexity via zero-communication protocols and applications. In *IEEE 53rd Annual Symp. on Foundations of Computer Science, New Brunswick, NJ, 20–23 October*, pp. 500–509. Piscataway, NJ: IEEE.

31. Touchette D. 2014 Quantum information complexity and amortized communication. (<http://arxiv.org/abs/1404.3733>)
32. Bravyi S, Hastings MB, Verstraete F. 2006 Lieb-Robinson bounds and the generation of correlations and topological quantum order. *Phys. Rev. Lett.* **97**, 050401. (doi:10.1103/PhysRevLett.97.050401)
33. Kitaev A, Preskill J. 2006 Topological entanglement entropy. *Phys. Rev. Lett.* **96**, 110404. (doi:10.1103/PhysRevLett.96.110404)
34. Levin M, Wen X-G. 2006 Detecting topological order in a ground state wave function. *Phys. Rev. Lett.* **96**, 110405. (doi:10.1103/PhysRevLett.96.110405)
35. Alicki R, Fannes M. 2004 Continuity of quantum conditional information. *J. Phys. A Math. Gen.* **37**, 55–57. (doi:10.1088/0305-4470/37/5/L01)
36. Müller-Lennert M, Dupuis F, Szehr O, Fehr S, Tomamichel M. 2013 On quantum Rényi entropies: a new generalization and some properties. *J. Math. Phys.* **54**, 122203. (doi:10.1063/1.4838856)
37. Wilde MM, Winter A, Yang D. 2014 Strong converse for the classical capacity of entanglement-breaking and Hadamard channels via a sandwiched Rényi relative entropy. *Commun. Math. Phys.* **331**, 593–622. (doi:10.1007/s00220-014-2122-x)
38. Fuchs CA. 1996 Distinguishability and accessible information in quantum theory. PhD thesis, University of New Mexico. (<http://arxiv.org/abs/quant-ph/9601020>)
39. Sion M. 1958 On general minimax theorems. *Pacific J. Math.* **8**, 171. (doi:10.2140/pjm.1958.8.171)
40. Junge M, Renner R, Sutter D, Wilde MM, Winter A. 2015 Universal recovery from a decrease of quantum relative entropy. (<http://arxiv.org/abs/1509.07127>)
41. Holevo AS. 2012 *Quantum systems, channels, information*. De Gruyter Studies in Mathematical Physics 16. Berlin, Germany: De Gruyter.
42. Bourbaki N. 1966 *Elements of mathematics: general topology*. Hermann, MO: Éditions des Sciences et des Arts.
43. Furrer F, Åberg J, Renner R. 2011 Min- and max-entropy in infinite dimensions. *Commun. Math. Phys.* **306**, 165–186. (doi:10.1007/s00220-011-1282-1)
44. Fuchs CA, van de Graaf J. 1999 Cryptographic distinguishability measures for quantum-mechanical states. *IEEE Trans. Inform. Theory* **45**, 1216–1227. (doi:10.1109/18.761271)
45. Tomamichel M, Colbeck R, Renner R. 2010 Duality between smooth min- and max-entropies. *IEEE Trans. Inform. Theory* **56**, 4674–4681. (doi:10.1109/TIT.2010.2054130)
46. Tomamichel M. 2012 A framework for non-asymptotic quantum information theory. PhD thesis, ETH Zurich. (<http://arxiv.org/abs/1203.2142>)
47. Wilde M. 2013 *Quantum information theory*. Cambridge, UK: Cambridge University Press.
48. Barnum H, Knill E. 2002 Reversing quantum dynamics with near-optimal quantum and classical fidelity. *J. Math. Phys.* **43**, 2097–2106. (doi:10.1063/1.1459754)
49. Nielsen MA, Chuang IL. 2000 *Quantum computation and quantum information*. Cambridge, UK: Cambridge University Press.
50. Sutter D, Tomamichel M, Harrow AW. 2015 Monotonicity of relative entropy via pinched Petz recovery map. (<http://arxiv.org/abs/1507.00303>)
51. Hirschman II. 1952 A convexity theorem for certain groups of transformations. *J. d'Analyse Math.* **2**, 209–218. (doi:10.1007/BF02825637)